

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24年 6月11日現在

機関番号：12601

研究種目：研究活動スタート支援

研究期間：2010～2011

課題番号：22800007

研究課題名（和文） 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に関する研究

研究課題名（英文） Skeletal Parallel Programming for Non-numerical Problems

研究代表者

江本 健斗 (EMOTO KENTO)

東京大学・大学院情報理工学系研究科・助教

研究者番号：00587470

研究成果の概要（和文）：

データ列上の非数値計算問題に対し、愚直なプログラムの容易かつ柔軟な記述とその記述からの効率的な並列プログラムの自動導出とを提供する理論と、その実現であるライブラリとを構築できた。これは、本研究で到達したかった目標点である。最適化規則の開発にはスケルトンプログラミングの研究で得た知見が応用されており、また、提唱する生成検査集約プログラミングは3種類のスケルトンで問題を構造化するという新しいスケルトン並列プログラミングの形を与えるものである。この意味において、本研究はスケルトン並列プログラミングを非数値計算問題に対して応用できたと言える。また、生成検査集約プログラミングは動的計画法アルゴリズムの改良・再利用のための枠組とも捉えることができ、アルゴリズムの系統的・機械的導出のための新しい手法を構築してきたとも考えられる。

研究成果の概要（英文）：

We have developed (1) a theory that provides an automatic optimization mechanism to derive efficient parallel programs from user programs naively-written in a concise notation for non-numerical problems on sequences, and (2) a library that provides the notation and the automatic optimization. This research has applied the skeletal parallel programming to non-numerical programs: The optimization has been developed based on the theory of skeletal parallel programming, and in our proposing GTA (Generate, Test, and Aggregate) programming a program is built by three kinds of components, each of which can be seen as a kind of skeletons. In addition, our GTA programming can be seen as a new programming framework for dynamic programming algorithms, in a point of view of program development/derivation.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2010年度	920,000	276,000	1,196,000
2011年度	1,160,000	348,000	1,508,000
年度			
年度			
年度			
総計	2,080,000	624,000	2,704,000

研究分野：総合領域

科研費の分科・細目：情報学・ソフトウェア

キーワード：並列プログラミング、生成検査集約法、非数値計算問題、組み合わせ問題、プログラム導出、半環、スケルトン、動的計画法

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、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。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。その代表的な問題は、グラフ上の最適化問題や動的計画法などの組み合わせ最適化問題であり、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単純な数値を扱う数値計算であり、それらの手法は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、非数値計算問題に対する実用的な並列プログラミングの枠組の構築を目指し、スケルトン並列プログラミングの立場から下に示す 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る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) 柔軟な記述性を持つ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、単純なリスト・行列・木などのデータ構造上の規則的な計算パターンを中心に定式化が行われていた。しかし、実際の組み合わせ最適化問題などでは、必ずしもそのような単純な規則的なパターンだけでは計算を捉えきることができない。例えば、多くの問題において、これらのパターンを複雑にネストさせる必要がある。そこで、非数値計算問題に必要な処理を適切に抽象化した新しい並列スケルトンの定式化を行い、非

数値計算問題を柔軟に記述できる並列スケルトンの集合を提案する。特に、ネストした計算を記述するためのデータ並列言語 NESL や新しい並列言語 Fortress にある生成子 (generator) の機能を参考に、柔軟な記述性をもつ定式化を与える。さらに、新しく定式化する並列スケルトンに対して、スケルトン並列プログラミングの強みであるプログラム最適化を提案する。

(2) 柔軟で且つ高い記述性をもつスケルトンライブラリの実現： 非数値計算問題にも有効な、実用的スケルトン並列プログラミングに必要な柔軟となる記述性・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、スケルトン並列プログラミングと関連の深い関数プログラミングの特長が、次期 C++ 標準の C++0x や Fortress などの新しい手続き型言語に積極的に取り入れられている。これらの技術を応用して、記述がしやすく、最適化能力の高い並列スケルトン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までできるだけ避けるようにする。

(1) 柔軟な記述性を持つ並列スケルトンの定式化： (1-1) 記述対象とする具体的問題例の選定、(1-2) 非数値計算問題記述並列スケルトンの定式化、(1-3) 新しい定式化における最適化ルールの開発

(2) 柔軟で且つ高い記述性をもつスケルトンライブラリの実現： (2-1) 定式化に基づく並列スケルトンの実装・評価、(2-2) 記述性・最適化機能に関する実装・評価

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

理論の観点では、まず、非数値計算問題をプログラミングの観点から整理し、関数型プログラミングで広く利用されてきた内包表記による生成検査 (generate-and-test) の記述に縮約演算を組み合わせ、最短路問題や最適列問合せなどの組み合わせ問題のプログ

ラムが統一的に記述された。次に、それらの多くの問題に共通して現れる数学的な構造（半環）に着目し、既存のスケルトンの組み合わせで表現される効率的な並列プログラムを統一的な記述から導出する手法が開発された。本手法は、検査述語ないし問合せ述語の構造を用いて半環を再帰的に拡張することを特徴とし、既存の導出手法にくらべ、（１）半環の構造を用いるため効率的な並列プログラムが導出できる、（２）より広い範囲の問題に対して統一的な効率化を与えられる、（３）半環という数学的に整理された概念を用いるため導出過程が理解しやすい、等の利点を持つ。また、本手法で用いるプログラム効率化のための変換規則は構成的に証明されており、その変換を機械的に用いることが可能である。そのため、本手法によるプログラムの効率化を自動化したプログラミング環境の構築が可能となる。

構築した基礎理論を元に、データ列上の組み合わせ問題等の非数値計算問題に対して、広く統一的に愚直なプログラム記述から効率的な並列プログラムを得ることのできる「生成検査集約プログラミング」の枠組を提唱した。この枠組は、（１）生成・検査・集約という３種類の部品に分けてプログラムを構成することによるプログラムの容易な設計、（２）逐次添加的に構成部品を追加・変更できることによるプログラムの容易な改善・拡張、（３）それら容易に記述できるプログラムからの最適化規則による効率の良いプログラムの導出、という特徴を持つ。これらの特徴を元に、ナップサック問題・最尤系列推定問題・部分列問合せ問題などに対し、愚直な記述から効率的な並列プログラムを機械的に得られること、過去の研究では個別に扱われていた拡張を統一的に扱えること、さらなる問題の拡張にも柔軟に対応できること、等を示した。また、提案された枠組は動的計画法アルゴリズムの新しい系統的開発の枠組とも捉えられ、この視点からの応用も期待される。

実践の観点では、簡単なプロトタイプ実装により、上記手法による最適化が機械的に行うこと及びその結果得られるスケルトンプログラムが正しく動作することが、具体的な例題を通して確認された。これにより、非数値計算問題のための最適化機能を備えたライブラリや領域特化言語が実現可能であることが示された。さらに、愚直な生成検査集約プログラムを記述するだけで効率的な MapReduce プログラムを得ることのできる最適化機能付きライブラリを Java と Hadoop の上で実現した。このライブラリを用いて上に挙げた具体的な問題に対する生成検査集約プログラムを記述・実行し、提案した枠組の有効性を確認した。

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 3 件）

- ① 江本 健斗：生成検査+ α 計算の効率的並列アルゴリズムの系統的導出. コンピュータソフトウェア, (査読有り), Vol. 29, pp. 159--175, 2012.
- ② Kento EMOTO, Sebastian FISCHER and Zhenjiang HU: Generate, Test, and Aggregate---A Calculation-based Framework for Systematic Parallel Programming with MapReduce---. Mathematical Engineering Technical Reports, (査読無し), 2011-34, pp. 1--28, 2011.
- ③ Kento EMOTO: An Algebraic Approach to Efficient Parallel Algorithms for Nested Reductions. Mathematical Engineering Technical Reports, (査読無し), 2011-01, pp. 1--31, 2011.

〔学会発表〕（計 5 件）

- ① Kento Emoto, Sebastian Ficsher, Zhenjiang Hu: Generate, Test, and Aggregate---A Calculation-based Framework for Systematic Parallel Programming with MapReduce---. 21st European Symposium on Programming (ESOP 2012), (査読有り), March 27th 2012, Tallinn, Estonia.
- ② Kento Emoto, Sebastian Ficsher, Zhenjiang Hu: Generate, Test, and Aggregate---A Calculation-based Framework for Systematic Parallel Programming with MapReduce---. プログラミング及びプログラミング言語ワークショップ PPL2012, (査読無し), 2012年3月9日, 和歌山県西牟婁郡白浜町.
- ③ 劉 雨, セバスティアン フィッシャー、江本 健斗、胡 振江：Generate-Test-and-Aggregate アルゴリズムの Hadoop 実装. 日本ソフトウェア科学会第 28 回大会, (査読無し), 2011年9月27日, 沖縄県那覇市.
- ④ Kento EMOTO: An Algebraic Approach to Efficient Parallel Algorithms for Nested Reductions. プログラミング及びプログラミング言語ワークショップ PPL2011, (査読有り), 2011年3月10日, 北海道札幌市.
- ⑤ Kento EMOTO: Semirings for Free! - An Algebraic Approach to Efficient Parallel Algorithms for Nested

Reductions. The Fourth DIKU-IST Joint
Workshop on Foundations of Software,
(査読無し), 13th January 2011,
University of Tokyo.
[図書] (計0件)

[産業財産権]

○出願状況 (計0件)

名称 :
発明者 :
権利者 :
種類 :
番号 :
出願年月日 :
国内外の別 :

○取得状況 (計0件)

名称 :
発明者 :
権利者 :
種類 :
番号 :
取得年月日 :
国内外の別 :

[その他]

ホームページ等

6. 研究組織

(1) 江本 健斗 (EMOTO KENTO)

東京大学・大学院情報理工学系研究科・助
教

研究者番号 : 00587470

(2) 研究分担者

()

研究者番号 :

(3) 連携研究者

()

研究者番号 :