

令和 6 年 6 月 10 日現在

機関番号：17104

研究種目：基盤研究(C)（一般）

研究期間：2019～2023

課題番号：19K11903

研究課題名（和文）正しさと効率の形式的証明を備えたスケルトン並列プログラミング環境に関する研究

研究課題名（英文）Research on Skeletal Parallel Programming Environment with Formal Proofs of Correctness and Efficiency

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・准教授

研究者番号：00587470

交付決定額（研究期間全体）：（直接経費） 3,300,000円

研究成果の概要（和文）：本研究では、定理証明支援系での形式証明を経た正しい（並列）プログラムの構築に対し、その支援技術の開発や具体的なモデル・言語機構の形式化を行った。形式化については、計算量を含めた低層並列計算モデル、並列計算のための負荷分散等を実現できる言語機能、並列プログラムの構成パーツ等の定式化の基礎となる recursion schemes の定理証明支援系での形式化を行った。支援技術については、これらの証明を単純化するためのライブラリやシステムの構築を行った。これらの取り組みで得られた成果は、国際会議や国内雑誌、国内大会・研究会等で発表した。

研究成果の学術的意義や社会的意義

昨今、複数のコンピュータや複数のCPU・コアを用いて計算を加速する並列計算は、AI・物理シミュレーション・ビッグデータ処理の実現において必須の技術となっている。しかし、並列計算を実現する並列プログラムの開発は難しく、特に正しく効率の良い並列プログラムが得られていることを保証することは困難である。この問題に対し、本研究は、定理証明支援系による正しさと効率の形式的保証をもった並列プログラムの体系的な開発環境を構築しようとするものであり、得られた研究成果はこの問題の解決に貢献する。

研究成果の概要（英文）：In this research, we developed formalization of parallel models and language systems on Coq and support tools for such formal proofs, to realize programming environments for certified correct/efficient parallel programs. This includes formalization of (1) the BSP model with formal proofs on user programs' computational costs, (2) new language systems useful for load-balancing in parallel programs, (3) recursion schemes that are the basis of parallel skeletons with powerful optimizations, as well as development of support tools such as libraries and semi-automatic techniques to reduce the burden of formal proofs. These results have been published in international conferences, domestic journals, and research workshops.

研究分野：計算機科学

キーワード：並列プログラミング 並列スケルトン 定理証明支援系 Coq 形式証明

1. 研究開始当初の背景

昨今、計算機で扱おうとするデータは巨大化の一途をたどり、また、計算機ハードウェアの性能向上は主にその並列性の向上によるところが大きい。そのため、現実的な時間での処理の実現には、ハードウェアの並列性を有効活用する並列プログラムの開発が必須となっている。しかし、並列プログラムの作成は、逐次プログラムの作成に比べて考慮すべきことが多く、プログラミング言語環境のサポートなしには困難である。並列プログラムの簡便な開発をサポートする並列プログラミング環境の一種として、スケルトン並列プログラミングの枠組みが知られている。スケルトン並列プログラミング環境では、ユーザは、並列スケルトンと呼ばれる並列計算パターンを与えられ、それに自身のアプリケーション特有の逐次コードを組み合わせることで目的の並列プログラムを構築する。並列性に起因する多種の面倒が並列スケルトンに隠蔽されるため、この枠組での並列プログラミングは非常に容易である。しかしながら、並列プログラミングにおける全ての面倒を解決するものではない。並列プログラムには、当然ながら、正しさと効率が求められる。正しくないプログラムには意味がなく、効率のよくないプログラムは望ましくない。開発した並列プログラムの正しさやその効率(並列計算量)を、そのプログラムの実行前に保証することが生産性の観点からは重要である。プログラムの正しさの検証には、最近、定理証明支援系(例えば Coq)での証明が用いられるようになってきた。また、プログラムの計算量についての証明手法も研究されるようになってきた。しかし、既存のスケルトン並列プログラミング環境は提供する並列スケルトンの定式化を与えないため、並列スケルトンを用いたユーザプログラムの正しさや計算量の証明は困難である。言い換えれば、既存の枠組みはユーザの書いたプログラムの正しさや効率に関しては無責任であった。

2. 研究の目的

本研究の目的は次の三つである。

(1) 開発・最適化・証明を一貫して行えるスケルトン並列プログラミング環境の構築。

本研究では、既存の環境が蔑ろにしていた「正しさの証明」や「計算量の証明」へのサポートを統合した、新しいスケルトン並列プログラミング環境の構築を目的の一つとした。

(2) 定理証明支援系における計算量証明等を目的とした証明手順簡略化手法の開発。

本研究では、プログラムの正しさや計算量の証明に頻出するパターンに対し、その証明を手短に、また、直観的に記述できるようにするための手法の開発を目的の一つとした。

(3) 並列実行モデルの定理証明支援系におけるモデル化と並列スケルトンへの適用。

本研究では、ユーザ並列プログラムの各並列実行環境における並列計算量を証明可能とするために、その実行基盤となるバルク同期並列(BSP)モデルやワークスティーリング等の、低レベル並列実行環境の定理証明支援系における形式化を目的の一つとした。

3. 研究の方法

上記の目的に沿って、それぞれ以下のように研究を進めた。

(1) 並列スケルトンの基盤となる数学的道具立てを定理証明支援系 Coq 上に形式化し、それを用いて既存の並列スケルトンとその最適化規則との形式化を Coq 上で行う。また、コード抽出によって証明・最適化後の実コードを得られる環境を構築する。

(2) 定理証明支援系における計算量証明や正しさの証明を整理し、そこで必要とされる証明手順を簡略化するライブラリを開発する。また、系統的な手法の適用が困難な場合には、深層学習等の手法を用いた証明の(半)自動化などの実用的な手法の開発に取り組む。

(3) 並列スケルトンの実行基盤となる並列実行環境として、計算量の取り扱いを考慮したモデルの定理証明支援系 Coq での定式化に取り組む。特に、よく使われる既存の並列スケルトンがその基盤としているバルク同期並列(BSP)モデルの定式化に取り組む。また、ワークスティーリング等の並列実行基盤を実現する言語機構についての形式化にも取り組む。

4. 研究成果

(1) 並列スケルトンの数学的な定式化の基礎である recursion schemes の定理証明支援系上での形式化を行った。map や reduce に代表される、効率化のための融合変換を備えたスケルトンは、数学的に recursion schemes により定式化されている。この定式化により、様々なデータ構造に対する並列スケルトンが統一的に扱えるようになってきている。本研究は、並列スケルトンに対する統一的な形式的検証を可能とするために、一般的な recursion schemes を定理証明支援系 Coq で形式化した。この研究成果は査読付き国際会議 APLAS2019 に採択され、発表した。

(2) 並列スケルトンの一種である Hadoop MapReduce を用いたアプリケーションとして、大

規模ランダムグラフの生成を行うプログラムを構築した。これは、大規模なデータを扱うため、想定通りの計算量であることが強く期待される並列プログラムである。これについては国内学会の大会にて口頭発表を行い、その後査読付き国内雑誌に採録された。

(3) 並列実行モデルのひとつである BSP モデルについて、ユーザプログラムの計算量証明を行える形で定理証明支援系での形式化を行った。この BSP モデルは、よく知られた map や reduce などの並列スケルトンをはじめ、近年大規模グラフの並列計算に応用される頂点主体計算の基盤となっている。本研究の形式化により、これらを用いて作られたユーザプログラムに対して計算量(と正しさ)の形式的証明を行うことが可能となる。実際に、いくつかの並列スケルトンとそれを用いたアプリケーション(最大部分和計算)について計算量の証明を伴う形式的証明を行った。ただし、その証明の構築は複雑であり、その点を支援する技術の開発の必要性が明らかになった。この成果は国内ワークショップで発表された。

(4) 頂点主体計算や部分グラフ主体計算を直接用いるよりもグラフ計算を記述しやすい、大規模グラフ並列計算用の領域特価言語について、そのグラフ並列スケルトン群等へのコンパイルの改良や記述性向上の改良等を行った。また、諸々のグラフ並列計算プログラムの性能評価に使うランダムグラフの並列生成に関する研究も進めた。これらの成果は国内ワークショップで発表された。

(5) 動的負荷分散による並列プログラム等を記述できる、計算状態操作機構の定理証明支援系 Coq 上での形式化を行った。計算状態操作機構とは、計算状態(関数呼び出し先祖のローカル変数など)の動的な変更を安全に行うための機構の総称であり、動的負荷分散による効率的な並列計算の実現に有用である。本研究では、計算状態操作機構のうちの持続型例外処理機構について、Scheme 言語をベースとして定理証明支援系 Coq 上にそのセマンティクス形式化を与えた。これにより、持続型例外処理を用いて記述されたプログラムについて、その仕様(望ましくない実行状態にならないことや、終了時に正しい計算結果が得られていることなど)を形式的に検証することが可能となった。この成果は国内ワークショップで発表された。

(6) 定理証明支援系による証明構築の困難さの解消に向けて、その一部を深層学習的なアプローチにより自動化する手法の試作・評価を行った。定理証明支援系での形式証明の構築の完全自動化は、理想的ではあるものの、既存研究の結果から(少なくとも現状では)現実的とはいえない。この理想を一段現実的にした手法として、「証明の中間状態の大雑把な形を与え、その形に至るまでの証明を自動で構築する」というシステムを、深層学習モデルを用いて構築し、その性能や使い勝手を評価した。結果として、まだまだモデルの改良が必要なものの、アプローチとしては有用であるとの結論が得られた。今後は新たな研究課題としてこの手法の完成を目指す。この成果は国内ワークショップで発表された。

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 1件/うち国際共著 0件/うちオープンアクセス 1件）

1. 著者名 神野 薫, 江本 健斗	4. 巻 37
2. 論文標題 Watts-Strogatz モデルに基づく大規模ランダムグラフの分散並列生成	5. 発行年 2020年
3. 雑誌名 コンピュータ ソフトウェア	6. 最初と最後の頁 34--45
掲載論文のDOI (デジタルオブジェクト識別子) 10.11309/jssst.37.2_34	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

〔学会発表〕 計8件（うち招待講演 0件/うち国際学会 1件）

1. 発表者名 鶴狩 慧久, 江本 健斗
2. 発表標題 指定した形の中間証明状態へ至る形式証明の自動生成を目指して ~ Transformer を用いた深層学習的アプローチ ~
3. 学会等名 火の国情報シンポジウム2024
4. 発表年 2024年

1. 発表者名 福島 央章, 江本 健斗
2. 発表標題 大域的グラフ計算記述言語の最適化および辺集合に関する拡張
3. 学会等名 火の国情報シンポジウム2023
4. 発表年 2023年

1. 発表者名 田中 匠海, 江本 健斗
2. 発表標題 並列計算量の形式的証明を伴う BSP プログラム用 Coq ライブラリ
3. 学会等名 火の国情報シンポジウム2022
4. 発表年 2022年

1. 発表者名 森 公哉, 江本 健斗
2. 発表標題 定理証明支援系Coqによる持続型例外処理機構の形式化
3. 学会等名 火の国情報シンポジウム2022
4. 発表年 2022年

1. 発表者名 Kosuke Murata, Kento Emoto
2. 発表標題 Recursion Schemes in Coq
3. 学会等名 Programming Languages and Systems - 17th Asian Symposium, APLAS 2019 (国際学会)
4. 発表年 2019年

1. 発表者名 村田 康佑, 江本 健斗
2. 発表標題 Coq における Hyломorphism を用いたプログラム演算の検証に向けて
3. 学会等名 第22回プログラミングおよびプログラミング言語ワークショップ (PPL 2020)
4. 発表年 2020年

1. 発表者名 神野 薫, 江本 健斗
2. 発表標題 Watts-Strogatz モデルに基づく大規模ランダムグラフの分散並列生成
3. 学会等名 日本ソフトウェア科学会第36回大会
4. 発表年 2019年

1. 発表者名 村田 康佑, 江本 健斗
2. 発表標題 高度な運算定理の Coq による証明とその自動化
3. 学会等名 日本ソフトウェア科学会第36回大会
4. 発表年 2019年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------