

令和 2 年 5 月 22 日現在

機関番号：12601

研究種目：若手研究(B)

研究期間：2015～2019

課題番号：15K15965

研究課題名（和文）代数的性質と型システムに基づく自動並列化

研究課題名（英文）Program Parallelization based on Algebraic Properties and Types

研究代表者

森畑 明昌（Moriyama, Akimasa）

東京大学・大学院総合文化研究科・准教授

研究者番号：10582257

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）：本研究の目的はリダクション並列化のための基礎理論を与えることである。このためには、現実的なプログラム中の複雑なリダクションを扱えるだけの表現力と、効率の良い並列プログラムの構築を補助する型システムが求められる。本研究では、プログラミング言語理論で最も普遍的に用いられるラムダ計算と呼ばれる体系に代数的性質に基づく式の単純化を加えることでリダクション並列化が自然に定式化できることを示し、さらに式の単純化が成功することを保証する型システムを与えた。

研究成果の学術的意義や社会的意義

近年CPUの単体での性能が頭打ちになりつつこと、また多くの計算機が自然にネットワークを介して協調動作できるようになったことにより、並列計算の重要性は非常に高まっている。しかし、どのようなプログラムであれば、コンパイラによって自動的に効率の良い並列プログラム変換できるのかについてをよく説明できる理論は確立していない。本研究では、データの総和や平均などを求める「リダクション」と呼ばれる並列計算パターンに着目し、これをプログラミング言語理論として自然に、しかも実際のプログラム並列化に素直に繋がる形で、扱うことのできる理論的基盤を与えた。

研究成果の概要（英文）：This research project aimed to develop a foundation for studying reduction parallelizations. The requirements for the foundation are the following two. First, it should be able to describe complex and realistic programs and reductions. Second, it should be equipped with a type system that supports efficient parallel reductions. In this research, we showed that a lambda calculus, which is the best-studied framework for studying general-purpose programming languages, extended with algebraic simplification can capture a variety of parallel reduction patterns; moreover, we provided a type system that guarantees successful algebraic simplification.

研究分野：プログラミング言語

キーワード：並列計算 リダクション並列化 ラムダ計算 型システム 代数的性質 関数型言語

様式 C - 19、F - 19 - 1、Z - 19 (共通)

1. 研究開始当初の背景

近年 CPU の単体での性能が頭打ちになりつつこと、また多くの計算機が自然にネットワークを介して協調動作できるようになったことにより、並列計算の重要性は非常に高まっている。しかし、どのようなプログラムであれば、コンパイラによって自動的に効率の良い並列プログラム変換できるのかをよく説明できる理論は確立していなかった。特に、データの総和や平均などを求める「リダクション」と呼ばれる並列計算パターンについての議論は不十分であった。そのため、現状の並列プログラミング環境は、限定的なパターン(例えば配列値の総和)について並列化できるような仕組みをアドホックに与えているに留まっていた。

一方、リダクション並列化が総和のような自明な場合以外にも可能であることは複数の研究によって示唆されていた。例えば、Xu ら (APLAS 2003) は複数の演算子を用いるようなプログラムについて、Matsuzaki ら (SPAA 2006) は木構造処理について、さらに Fisher ら (PLDI 1994) や Sato ら (PLDI 2011) は条件分岐を伴うようなループについて、リダクション並列化が可能な場合があることを論じている。これらの結果は、リダクション並列化が実はかなり広い範囲のプログラムに対して適用可能であることを示唆する。しかし、これらの成果を統一的に説明できる理論は存在していなかった。そのため、これらの成果のさらなる一般化や、これらの成果に通底する一般的な着想の抽出などは困難であった。

2. 研究の目的

本研究の目的は、リダクション並列化を議論するための一般的で統一的な基礎理論を確立することである。この理論には、現実的なプログラム中の複雑なリダクションを扱えるだけの表現力と、リダクション並列化が実際に意味のある並列化となっているかの確認を補助するための型システムが求められる。この理論を確立することにより、リダクション並列化がどの程度広い範囲のプログラムに対して適用可能であるかを議論する。

3. 研究の方法

本研究では、プログラミング言語理論で最も普遍的に用いられる体系であるラムダ計算に基づき、標準的なラムダ計算にどのような構成要素を加えることでリダクション並列化を自然に扱えるようになるかを模索する。

リダクション並列化の文脈では、結合法則や分配法則などの代数的性質が重要であり、これらが欠けると並列化に失敗することがよく知られている。これをふまえ、プログラム中での演算子の使われ方とその代数的性質を追跡し、それによってリダクション並列化が成功するかどうかの判定を補助する型システムを与える。

4. 研究成果

本研究ではまず、代数的性質を用いた式の単純化の機能をもつラムダ計算を構築し、この体系の性質を調べた。この体系では「式の単純化」を通常の計算と並列に行うことでリダクション並列化の並列性を表現する。本研究では、本研究では「式の単純化」を行っても計算結果が変わらない、つまりこのアプローチによる並列化がプログラムの意味を変えないことを証明した。さらに、演算子の使われ方を追跡する型システムを与え、この型システムにより型が付くプログラムについては、「式の単純化」が少なくとも有害ではないこと、つまり単純化によって計算コストが著しく悪化したりしないことを保証した。さらにこの体系の表現能力について調査した。その結果、既存のリダクション並列化の文脈で議論されてきたほとんどの例に加え、既存手法では扱いが困難だと思われていたような例、例えば条件分岐を伴うジャンプや、各部分構造毎の集約結果を求める複数のアルゴリズムの比較など、についても比較的明快に扱うことができることが確認された。

以上の結果から、既存の高度なリダクション並列化手法に対して統一的な記述を与えることに成功したと言える。さらに、既存のリダクション並列化手法に通底する物が「式の単純化」であること、逆に「式の単純化」さえあれば広い範囲のプログラムに対してリダクション並列化が可能であることも明らかにしている。これは、より高度なリダクション並列化手法を論じる上での重要な基盤となると期待されれらる。

以上の成果の萌芽的な内容は第 18 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にて論文賞を受賞した。また、最終的な成果は、プログラミング言語分野でのトップ会議である ICFP (ACM International Conference on Functional Programming) 2019 に採択された。以上のように、この成果は国内外から高い評価を受けている。

また、主要な成果に加え、この研究の過程で得られた並列計算・リダクション並列化に関する知見からいくつか追加の成果が得られている。特に主要なものとしては、大規模グラフ並列分散処理に関する成果が挙げられる。大規模グラフの並列分散処理は、Web やソーシャルネットワーク等の解析のために必須だが、そのためのプログラムの系統的な構成法についてはまだ十分に理解されているとは言いがたい。これに対し、特に大規模グラフからの情報集約について、その情報集約に用いる演算子の性質によって効率を大きく改善できること、またそのような効率改

善は既存の制約解消機を用いることで自動化ができることを明らかにした。これにより、性能を意識せず記述した自明なグラフ処理プログラムから、効率の良い分散並列処理プログラムを自動的に得ることができる可能性が拓かれた。この成果は、プログラミング言語分野の主要な国際会議の一つである FLOPS (International Symposium on Functional and Logic Programming) 2018 に採択された。

5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 2件/うち国際共著 0件/うちオープンアクセス 0件）

1. 著者名 Moriyama Akimasa	4. 巻 27
2. 論文標題 Using Algebraic Properties and Function Fusion to Evaluate Tree Accumulations in Parallel	5. 発行年 2019年
3. 雑誌名 Journal of Information Processing	6. 最初と最後の頁 411 ~ 421
掲載論文のDOI (デジタルオブジェクト識別子) https://doi.org/10.2197/ipsjjip.27.411	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Moriyama Akimasa, Emoto Kento, Matsuzaki Kiminori, Hu Zhenjiang, Iwasaki Hideya	4. 巻 10818
2. 論文標題 Optimizing Declarative Parallel Distributed Graph Processing by Using Constraint Solvers	5. 発行年 2018年
3. 雑誌名 Lecture Notes in Computer Science	6. 最初と最後の頁 166 ~ 181
掲載論文のDOI (デジタルオブジェクト識別子) https://doi.org/10.1007/978-3-319-90686-7_11	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計7件（うち招待講演 1件/うち国際学会 1件）

1. 発表者名 森畑明昌
2. 発表標題 累積引数を伴う木構造走査の代数的性質と融合変換に基づく並列計算
3. 学会等名 情報処理学会プログラミング研究会第121回プログラミング研究発表会
4. 発表年 2018年

1. 発表者名 森畑明昌, 江本健斗, 松崎公紀, 胡振江, 岩崎英哉
2. 発表標題 頂点主体並列グラフ処理の制約解消器による効率化
3. 学会等名 日本ソフトウェア科学会第34回大会
4. 発表年 2017年

1. 発表者名 Akimasa Morihata
2. 発表標題 From Identification of Parallelizability to Derivation of Parallelizable Codes
3. 学会等名 5th ACM SIGPLAN Workshop on Functional High-Performance Computing (招待講演) (国際学会)
4. 発表年 2016年

1. 発表者名 Akimasa Morihata
2. 発表標題 Compilation from Lambda Calculus with Algebraic Simplification to Record Calculus
3. 学会等名 情報処理学会 第111回プログラミング研究発表会
4. 発表年 2016年

1. 発表者名 森畑明昌
2. 発表標題 頂点主体グラフ処理の構成的アルゴリズム論に基づく定式化
3. 学会等名 日本ソフトウェア科学会第33回大会
4. 発表年 2016年

1. 発表者名 森畑 明昌
2. 発表標題 代数的性質に基づくラムダ式の並列評価
3. 学会等名 第18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
4. 発表年 2016年

1. 発表者名 森畑 明昌
2. 発表標題 最適部分列問題に対する統一的な並列アルゴリズム
3. 学会等名 日本ソフトウェア科学会第29回大会
4. 発表年 2015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
--	---------------------------	-----------------------	----