

平成22年 5 月 17 日現在

研究種目：若手研究（B）
研究期間：2007～2009
課題番号：19700006
研究課題名（和文） 関数型プログラムの限界を越える安全なポインタを持つ再帰データ型の理論
研究課題名（英文） A Theory of Inductive Data Types Involving Safe Pointers
研究代表者
浜名誠（HAMANA MAKOTO）
群馬大学・大学院工学研究科・助教
研究者番号：90334135

研究成果の概要（和文）：

コンピュータ上での効率のよい計算に頻出する、サイクルと共有の構造を持つデータ構造を数理的構造を持つデータ型、すなわち再起データ型で表現する手法を探求し、関数型プログラミング技術への応用を行った。

研究成果の概要（英文）：

Cyclic sharing tree structures appear very often in logic and theoretical computer science. We investigate a mathematically clean representation of them. We apply the obtained representation to inductive datatypes in functional programming languages.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2007年度	600000	0	600000
2008年度	700000	210000	910000
2009年度	600000	180000	780000
年度			
年度			
総計	1900000	390000	2290000

研究分野：理論計算幾何学

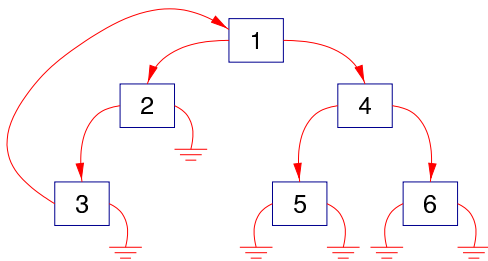
科研費の分科・細目：情報学・情報学基礎

キーワード：関数型プログラミング言語、項書換え・データ構造・情報基礎

1. 研究開始当初の背景

計算機中のデータ構造として、線形リストや木構造といったデータ構造は最も基本的なもので、どのようなソフトウェアの構築においても使われている。ソフトウェア基礎の観点では特に、線形リストは Lisp に代表される関数型言語の基本データ構造であり、木構造は、近年 XML データとして標準化され用いられている。

線形リストや木構造といった規則的なデータ構造は、数理的には再帰的な構造を持つものとして捉えられる。すなわちリストの先頭の次要素以降はまたリストを成し、木の部分構造はまた木を成すという構造である。このような再帰的なデータ構造を関数型言語では再帰データ型の機構により自然に定義し、パターンマッチの機構により直接的に操作することができる。このような関数型言語の明解さが C 言語や Java 言語に代表される手続き型言語に対する優位性であるが、一方再帰データ型で表すことができないデータ構造に対しては、関数型言語は不自然で非効率なプログラムになるという問題点がある。再帰データ型で表すことができないデータ構造の例として、以下のような巡回するノードを持つような木構造がある。



再帰的なデータ構造は、部分構造が必ず元の構造よりも小さいものでなくてはならないため、このような巡回している構造を再帰データ型で直接的に表すことは不可能と考えられている。手続き型言語では、ポインタの機構を用いてこのような巡回や共有構造を直接表すことができるが、関数型言語ではポインタのような低レベルの機構は、関数のような数学的概念とマッチしないため、直接扱うことができない。したがって、ポインタが本質的に必要なこのような構造、つまりグラフのようなデータ構造を関数型言語はうまく扱うことができなかった。グラフは多くの重要なアルゴリ

ズムで必須のデータ構造であるので、この欠点は関数型言語の実用性を限定する要因となっていた。

2. 研究の目的

申請者はこの問題に対する準備段階として、巡回構造を持つデータは、ネストデータ型というパラメータを持った再帰データ型を用いると表現可能となる事を示した研究を行った。但しさらに一般化の余地と、共有の構造は考察していないため完全な形でのポインタを持つ構造の定式化は未解決であった。本研究はこれを出発点として、次の二つを目的とした。

- (1) ポインタを直接用いることができないという関数型言語のデータ型表現の限界を、新たにポインタの代数的構造を明らかにすることにより解決する。
- (2) ポインタを持つデータ構造を用いるアルゴリズムを関数型言語の計算モデルである高階書換え系の体系で定式化し、その停止性を理論的に保証する手法を与える。

(1) については、データ型の表現に用いるためのポインタの代数構造を発見し、それを再帰データ型として表現するというアプローチを取る。これによって理論的明解さを壊すことなく関数型言語でのポインタを持つデータ構造を実現する。(2) については、ポインタを持つデータ構造は巡回を持つため、このデータを用いたプログラムは停止性しない危険性がある。停止性を事前に理論的に保証するために、項書換え系の研究分野での申請者のこれまでの研究結果を進展させて用い、ポインタを用いるプログラムの安全性を保証する。

3. 研究の方法

本研究は、関数型言語でのポインタを持つ再帰データ型の理論の構築と実装と、それを用いたアルゴリズムの停止性保証技術の確立を目指した。以下に、各項目の研究方法を区分して述べる。

[項目 1] ポインタを持つ再帰データ型の理論の構築と実装

ポインタを持つ再帰データ型の理論を構築し、それを関数型言語に実装して有効性を確認する。申請者の先行研究により、インデックス付き集合の圏の代数構造を用いて、高階抽象構文、高階書換え系、巡回を持つデータ型といった様々な体系が、インデックスを適切に変更することによってモデル化できることが分かっている。このアイデアを拡張し、先のインデックス集合として取られていた「変数」の集合を「ポインタ」の集合として捉え直し、ポインタの集合によってインデックス付けられた集合の圏、すなわち前層の圏の中で再帰データ型をモデル化するアプローチを取る。ここで最も重要な問題は、ポインタ集合の代数構造の厳密な構成方法である。本研究は安全なプログラムを構成することを目指すため、ここでのポインタは不正な領域を指さずに再帰データ型によって構成されたデータのみを指し示すものとして構成しなければならない。

次に、得られた再帰データ型の理論を関数型言語 Haskell に実装して有効性を確認する。Haskell は現存する関数型言語の中で最も先進的な型システムを持つことで知られている。それを用いることでインデックスされた集合としてモデル化したポインタ付き再帰型のような代数的構造を理論的に正しく実装する。

[項目 2] 高階書換え系の停止性保証技術の確立

高階書換え系による定式化、および高階書換え系の停止性技術の拡張と適用を行う。

高階書換え系の停止性技術を発展させたものを用いて、ポインタを持つ再帰データ型を用いるアルゴリズムの停止性を理論的に保証する理論を確立する。

これは与えられた高階項書換え系の停止性判定問題を考察することに他ならない。先に得られたポインタを持つ再帰データの代数構造を最大限に活用し、一階項書換え系の意味ラベル技術を高階に拡張する。代数的意味論を用いると、高階書換えの対象の高階項にラベル付けする操作をラベル付き項の代数的構造を考察することによって、自

然に準同型写像として定式化できるという利点があるため、円滑に拡張が行える。

意味ラベル技術の元となる関数型プログラムの意味論研究は、英国ケンブリッジ大学計算機研究所の Marcelo Fiore 博士による既存研究との深い関連がある。Fiore 博士とは、高階抽象構文に関する研究で互いに参照しあっており、本テーマの問題意識と知識を共有しているため、同博士との継続的な議論は、本研究の推進に極めて有用である。このために海外旅費を用いた研究訪問を行う。

4. 研究成果

(I) 高階書換え系の停止性技術を発展させたものを用いて、ポインタを持つ再帰データ型を用いるアルゴリズムの停止性を理論的に保証する技術を確立した。

これは与えられた高階項書換え系の停止性判定問題を考察することに他ならず、既存の高階項書換え系の停止性判定技法のみでは、不十分であることが分かっていた。そこで先に得られたポインタを持つ再帰データの代数構造を最大限に活用した。すなわち、申請者の先行研究の高階構文の代数的意味論の構造を用い、一階項書換え系の意味ラベル技術を高階に拡張した。代数的意味論を用いると、高階書換えの対象の高階項にラベル付けする操作をラベル付き項の代数的構造を考察することによって、自然に準同型写像として定式化できるという利点があるため、円滑に拡張が行えた。さらに General Schema と呼ばれる停止性判定技法とこの意味ラベルを組み合わせることで、停止性を証明する技法を確立した。

さらにこれに加えて高階書換え系についての新たな結果も得た。高階書換え系にソリッド性という性質を導入し、これに関する停止性のモジュラ性を意味ラベリング手法を用いて証明することに成功した。それと共に高階書換え系と高階プログラムスキームの合成に関する停止性のモジュラ性も証明することができた。この結果は国際会議 ACM-SIGPLAN International Conference on Principles and Practice of Declarative Programming で論文と共に発表した。

(II) ポインタを持つ再帰データ型のための始代数

理論を構築し、それを関数型言語に実装して有効性を確認することを行った。

これまでの研究からインデックス付き集合の圏の代数構造を用いて、高階抽象構文、高階書換え系、サイクルを持つデータ型といった様々な体系が、インデックスを適切に変更することによってモデル化できることが分かっている。本研究はこのアイデアを拡張し、先のインデックス集合として取られていた「変数」の集合を「ポインタ」の集合として捉え直し、ポインタの集合によってインデックス付けられた集合の圏で再帰データ型をモデル化するアプローチを取った。このために導入したアイデアはインデックスとして「構成するデータ型の形状」を取った点である。例えば、二分木型にサイクルと循環のポインタを導入するためには、木のノードの情報を取り除いた「木の形状」のみを持つ構造をインデックスとする。不正なポインタのないデータ型を構成することができた。

さらにこの理論によって得られたデータ型を関数型言語 Haskell に実装を行った。Haskell の先進的な型機構を最大限に活用した。この目的のために「一般化代数データ型 (GADT)」が適切であることを見出し、形状でインデックスされたデータ型を実現できた。

また執筆した論文「高階書換え系の停止性のための代数モデル」が日本ソフトウェア科学会第 12 回論文賞を受賞した。これは、長く未解決だった高階書換えの体系コンビナトリ-簡約系の完全な代数モデルを与え、メタプログラミングシステムのための計算体系としてメタ書換えの完全な代数的な特徴付けの結果を得たことが評価された。

(III) ポインタを持つ再帰データ型のための始代数理論を完成させ、プログラミングや推論に有用な様々な原理を導出した。

関数型言語において、始代数は特にパターンマッチと構造再帰、そして推論における構造帰納法と密接に関係している。ポインタを持つ再帰データ型に対して、この関係を詳細に調べ、具体的にそれらの原理がどのような形になるかを明らかにした。さらにポインタを持つ再帰データ型の依存型表現を依存型プログラミング言語 Agda を用いて明らかにした。この結果は Ninth International Conference on Typed Lambda Calculi and

Applications (TLCA 2009) で発表し、論文は Springer-Verlag から出版された。さらにこれを拡張したジャーナル版論文も学術論文誌 Logical Methods in Computer Science に採録決定している。

また、高階書換え系の停止性技術を発展させたものを用いて、ポインタを持つ再帰データ型を用いるアルゴリズムの停止性を理論的に保証する技術を確立した。先に得られたポインタを持つ再帰データの代数構造を最大限に活用し、一階項書換え系の意味ラベル技術を高階に拡張した。さらに高階の再帰経路順序の技法とこの意味ラベルを組み合わせることにより、停止性を証明する技法を確立することができた。この結果は 18th International Workshop on Functional and (Constraint) Logic Programming で発表し、論文は Springer-Verlag から出版された。

5. 主な発表論文等 [雑誌論文](計 6 件)

- [1] [M. Hamana](#). Initial Algebra Semantics for Cyclic Sharing Tree Structures. Logical Methods in Computer Science, 掲載決定, 査読有.
- [2] [M. Hamana](#). Semantic Labelling for Proving Termination of Combinatory Reduction Systems, In Proc. of 18th International Workshop on Functional and (Constraint) Logic Programming, Lecture Notes in Computer Science 5979, p.62-78, Springer-Verlag, 2010, 査読有.
- [3] [M. Hamana](#). Initial Algebra Semantics for Cyclic Sharing Structures, Typed Lambda Calculi and Applications, Lecture Notes in Computer Science 5608, pp.127-141, Springer-Verlag, 2009, 査読有.
- [4] 松田一孝, 胡振江, 中野圭介, 浜名誠, 武市正人. 補関数の生成による複製機能付きプログラムの自動双方向化, コンピュータソフトウェア, Vol. 26, No. 2, pp. 56-75, 岩波書店, 2009, 査読有.

- [5] M. Hamana. Higher-Order Semantic Labelling for Inductive Datatype Systems. In Proc. of Ninth ACM SIGPLAN International Conference on Principles and Practice of Declarative Programming, pp. 97-108, ACM Press, 2007, 査読有.
- [6] K. Matsuda, Z. Hu, K. Nakano, M. Hamana, and M. Takeichi. Bidirectionalization Transformation based on Automatic Derivation of View Complement Functions, In Proc. of 12th ACM SIGPLAN International Conference on Functional Programming, pp. 47-58, ACM Press, 2007, 査読有.

[学会発表](計 8 件)

- [1] 浜名誠, Inductive Cyclic Sharing Data Structures, 相互作用の幾何、トレース付きモノイダル圏と非明示的計算量研究会, 京都大学, 25 August, 2009.
- [2] 浜名誠, Initial Algebra Semantics for Cyclic Sharing Structures, Typed Lambda Calculi and Applications, 2009.7.3.
- [3] 浜名誠, Semantic Labelling for Proving Termination of Combinatory Reduction Systems, 18th International Workshop on Functional and (Constraint) Logic Programming, 2009.6.28.
- [4] 浜名誠, Inductive Cyclic Sharing Data Structures, Workshop on Mathematics for Pressing Problems in Computer Science, 京都大学, 30 June, 2008.
- [5] 浜名誠, Term Syntax for Cyclic Sharing Structures, 29th TRS meeting, 東京大学, 28 February, 2008.
- [6] 浜名誠, Representing a Cyclic Sharing Structure as a GADT, IPL seminar, 東京大学, 30 October, 2007.
- [7] 浜名誠, What is the Category for Haskell?, CAPS, IPL, 東京大学, 23 October, 2007.

- [8] 浜名誠, Higher-Order Semantic Labelling for Inductive Datatype Systems. Ninth ACM SIGPLAN International Conference on Principles and Practice of Declarative Programming, 2007.7.14.

[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

<http://www.cs.gunma-u.ac.jp/~hamana/>

6. 研究組織

(1) 研究代表者

浜名誠 (HAMANA MAKOTO)

群馬大学・大学院工学研究科・助教

研究者番号: 90334135