科学研究費助成事業 研究成果報告書

令和 4 年 6 月 3 日現在

機関番号: 11301

研究種目: 基盤研究(C)(一般)

研究期間: 2017~2021

課題番号: 17K00007

研究課題名(和文)定式化された形式木言語理論に基づくソフトウェア基盤技術の開発

研究課題名(英文)Development of software foundation based on certified formal tree language theory

研究代表者

中野 圭介 (Nakano, Keisuke)

東北大学・電気通信研究所・教授

研究者番号:30505839

交付決定額(研究期間全体):(直接経費) 3,400,000円

研究成果の概要(和文):形式木言語理論とは,形式的な木構造データを扱うことのできる木オートマトンや木トランスデューサに関する理論であり,効率的かつ検証可能なソフトウェアへの応用が期待されている.本研究期間ではいくつかの木トランスデューサのクラス間の関係について研究を進めたが,特筆すべき結果として,効率のよいとされるストリーム型木トランスデューサのクラスが,等価性判定問題が決定可能である最大のクラスと一致することを示すことに成功した.この研究成果により,ある種の効率的なプログラムの検証アルゴリズムの開発が期待される.

研究成果の学術的意義や社会的意義 二つのプログラムの振舞いが全く同じであるか(等価性)を検証する問題は一般に決定不能であるが,形式を制 限することで決定可能になることが知られている.この制限は非常に強く,特に効率を意識したプログラムにお いてはその複雑さから等価性を判定することは困難であるとされてきた.本研究の成果では,ストリーム型変換 と呼ばれる効率的かつ煩雑なプログラムについても等価性判定が決定可能であることが証明され,これはプログ ラム検証における新たな可能性を示している.提案した等価性判定の手法は,特定の条件下での等価性(部分等 価性)の判定も扱うことができるため,後方互換性を意識したソフトウェアの保守にも応用が可能である.

研究成果の概要(英文): Formal tree language theory is a theory of tree automata and tree transducers for a formal model of tree-structured data, and is expected to be applied to efficient and verifiable software. In this research project, we have studied the relation between several classes of tree transducers and succeeded in showing that the class of stream-type tree transducers that is considered efficient corresponds to the largest class in which the equivalence problem is decidable. The results of this research are expected to lead to the development of some efficient program verification algorithms.

研究分野: 形式木言語理論

キーワード: 形式木言語理論 プログラム検証 定理証明支援系 双方向変換

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等に ついては、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

1.研究開始当初の背景

本研究の重要な背景となる定理証明支援系と形式木言語理論に関する概要を研究代表者の研究背景と共に示したうえで,ソフトウェア検証における形式木言語理論の応用に関して簡単に述べる

定理証明支援系(proof assistants または interactive theorem provers)は,数学的な定理を計算機上で対話的に証明するためのシステムのことで,いくつかの実装やそのアイデア自体は古くから存在していたが,Coq や Agda や Isabelle/HOL や Lean などの定理証明支援系の普及とハードウェアの性能向上およびアルゴリズムの発展により,ここ数年で大きな注目を集めている技術の一つである.定理証明支援系の上で証明された定理は,一種のプログラムファイルとして証明と共に保存することができるため,他の利用者と共有したり再利用したりすることも可能である.研究代表者は平成 28 年度までの研究課題において,定理証明支援系を用いた形式木言語理論の定式化を進めてきた.

本研究課題において重要な基盤となる形式木言語理論(formal tree language theory)は,形式言語理論の一分野で,木構造をもつデータを形式的に扱うものであり,木オートマトンや木トランスデューサなどを対象とした理論である.形式木言語理論は,数学的な興味にとどまらず,データベース研究者の間では XML などの木構造データを処理するプログラムの正当性を検査するために利用され,プログラミング言語研究者の間では効率的なプログラムの自動導出やソフトウェアのモデル検査のための道具として用いられるなど,様々な場面で応用されている.本研究代表者は,平成 13 年頃から木トランスデューサの理論や応用に関する研究に取り組み,平成 24 年には木オートマトンと木トランスデューサの動向に関する国際ワークショップ(Trends in Tree Automata and Tree Transducers, TTATT)のプログラム共同委員長を務め,その後 3 回にわたる同ワークショップの全てにおいてプログラム委員を務めるなど,この分野において大きな関心をもって研究を進めており,今回の目的である形式木言語理論の定式化と応用に対して強い動機をもって取り組んできた.

形式木言語理論はソフトウェア検証において重要な役割を果たすことが期待されていながら, 実用的なプログラムに応用するには隔たりが大きく,広く利用されるまでには至っていない.

2.研究の目的

本研究の目的は,定理証明支援系において定式化された形式木言語理論を,ソフトウェア開発やプログラム検証へ応用するために拡張し,理論と実用の両面における研究基盤を確立することである.定理証明支援系は数学的な定理を計算機上で対話的に証明するためのシステムであり,このシステム上での数学理論の定式化が急速に進められている.本研究代表者は,平成 28年度までの研究課題において定理証明支援系を用いた形式木言語理論の定式化を進めてきたが,この定式化は教科書的なものにとどまっており,実際のソフトウェア開発・検証への応用には大きな隔たりがある.そこで,本研究課題では,実用に合わせた形式木言語理論の再定式化を行ったうえでその有用性を実証する.なお,定理証明支援系による定式化は,マイクロソフトの研究グループが数年間にわたり一つの群論の定理に証明を与えたことからもわかるように,上述の形式木言語理論の定式化作業は大変なものであると予想されていた.本研究では,平成 28年度までに行ってきた研究課題における定式化の経験を生かし,ここ数年で飛躍的に進歩した定理証明支援系 Coq の証明ライブラリ mathcomp を利用することで目標を達成する.

3.研究の方法

本研究では、研究目的で示した定理証明支援系によって定式化された形式木言語理論の実用化に向けて、次の3つの方針を軸に目的を達成することを目指していた。1つめはプログラム変換や合成などのソフトウェア開発に必要な形式木言語理論の定式化、2つめはモデル検査をはじめとするプログラム検証に必要な形式木言語理論の定式化、3つめは定理証明支援系のプログラム抽出機構や文書化機能を利用した実用基盤の公開である。1つめおよび2つめについては、前年度までの研究代表者の研究成果を拡張するが、必要に応じて関連する理論の定式化も行うことも考えられる。3つめは定理証明支援系Cogによって定式化を行うことで実現を試みた。

4. 研究成果

先述の通り,マイクロソフトの研究グループによる代数学の定式化に大きな困難が伴ったように,形式木言語理論の定式化においても応用に至るまでには期間内に終えられないものと期間中に判断し,定式化された形式木言語理論にこだわらず,ソフトウェア検証に応用可能な形式木言語理論の発展に注力した.結果的には,ソフトウェア検証に直接応用可能な理論の確立にも成功し,本研究の目的は果たせたといえる.以下では,各年度の研究成果について報告を行う.

[平成 29 年度]

前年度までの研究代表者の研究成果である木言語理論および木トランスデューサ理論の定式 化を整理し,実際の応用に必要な拡張について検討を行うのが本年度の目標であった.ここで具 体的に想定されていた応用とは,形式木言語理論を利用した (1)ソフトウェア開発の手法 (プログラム合成やプログラム変換) および (2)プログラム検証の手法 (モデル検査や等価性判定)のことである.本年度は,(1)としてスタック属性付き木トランスデューサの表現力の再考,(2)としてはプログラム検証である (2a) 木トランスデューサの型検査や (2b) 等価性判定の実装の準備を進めた.

- (1)については研究代表者が平成 21 年に発表したスタック属性付き木トランスデューサを再考し,プログラム変換への応用を発表したが,循環性の考慮が必要なその意味論の複雑さから定理証明支援系において定式化することは難しく,今後の課題であると言える.
- (2a) 木トランスデューサの型検査は,与えられた木変換規則と入出力の仕様に対し,「仕様を満たす入力を変換すると必ず仕様を満たす出力が得られる」ということを静的に検査することであり,逆像に基づく計算手法が一般的である.本研究では木オートマトンの表現を工夫することで効率的な実装ができることを確認した.
- (2b) 木トランスデューサの等価性判定は,与えられた2つの木変換が同じ入力に対して同じ出力を返すことを静的に判定することであり,降下型木トランスデューサについては様々なアルゴリズムが提案されているものの実装されているものはほとんどない.一部のアルゴリズムについて実装を進めている.

[平成 30 年度]

- 2年目にあたる平成30年度では,形式木言語理論の中心となる木トランスデューサ(木変換の形式的表現)と木オートマトン(木集合の形式的表現)の双方について応用やそのための拡張について研究を進めた.本年度の成果は,主に(1)木トランスデューサの記号的拡張,(2)木トランスデューサの等価性判定,(3)結合子論理の性質の部分的解明である.
- (1)の記号的拡張は既存の木オートマトンや木トランスデューサを拡張する手法の一つであり 記号の有限性制約を外すことを目的としている.この拡張は,有限性を満たす場合の望ましい性 質を維持しながら注意深く行う必要があるが,本研究では,マクロ木トランスデューサとよばれ る比較的表現力の高いものについて拡張に成功した.
- (2)の等価性判定はプログラム検証の手法に応用可能な「同じ入力に対し必ず同じ出力が得られる」という問題を解くことであり、一般的には決定不能問題であるが、特定の木トランスデューサについては決定可能であることが知られている。本研究では、ストリーム木トランスデューサの特殊な場合について等価性が決定可能になると予想し、部分的に解決した。
- (3)の結合子論理の中で関数合成を表す結合子について,研究代表者が平成 20 年に予想した性質を部分的に証明した.この証明においては,無限個の要素を含む木集合を有限で表せる木オートマトンを利用している.

[令和元年度]

- 3年目にあたる令和元年度では,木構造データ変換の形式モデルである木トランスデューサについて研究を進めた.主な成果は,研究代表者がこれまでに研究を進めてきたスタックを用いた木トランスデューサの延長となるものであり,(1)木トランスデューサの等価性判定に関する結果と(2)既存の木トランスデューサのクラスに対するスタック機構の導入である.
- (1)の木トランスデューサの等価性判定についての結果は、ここ数年で明らかになった「文字列から木への下降型木変換器」に対する等価性判定アルゴリズムを応用するものである.具体的には、スタックを用いた効率のよい木構造処理を直接表現した木変換のクラス「ストリーム型木トランスデューサ」を扱い、スタックを扱うクラスのトランスデューサのうち、等価性判定のアルゴリズムが存在する最も汎用的なクラスの発見に成功した.
- (2)の木トランスデューサのクラスの拡張においては,研究代表者が提案した属性付き木トランスデューサに対するスタック機構の導入を応用し,マクロ木トランスデューサと呼ばれる上位のクラスに対して同様の拡張を行った.本研究により,スタック機構の導入した場合には,この下位と上位のクラス関係が必ずしも成立しなくなる可能性があることを確認した.

この他の成果として,研究代表者のテーマの一つであった木構造データに対する双方向変換についての研究も並行して進めた.双方向変換は,二つのデータ間を同期するために必要な二つの変換のことであるが,それらの一貫性を保証するために必要な複数の双方向変換則に対し,これまで知られていなかった相互の関係を発見した.

[令和2年度]

最終年度にあたる令和 2 年度では,形式木言語理論における木構造データ変換のモデルであるストリーム型の木トランスデューサに着目して研究を進めた.主な成果は,研究代表者がこれまでに研究を進めてきた木トランスデューサの効率的な実装の延長となるものであり,(1)ストリーム型木トランスデューサの表現力の特定と(2)等価性判定が決定可能な木トランスデューサに関する結果である.

(1)のストリーム型木トランスデューサの表現力については、比較的自然な制約の下で既知の木トランスデューサのクラスと一致することを証明した、ストリーム型トランスデューサについては、研究代表者が平成16年に木トランスデューサ理論を構造化文書処理の効率化に応用して以来形式的な表現力が未解決であったが、今回の研究成果によりこの問題に対する部分的な

解決が得られたと言える.

(2)の等価性判定が決定可能な木トランスデューサに関する結果についても,ストリーム型木トランスデューサのクラスに対して研究を行った.木トランスデューサの等価性判定に関する従来の結果の多くはストリーム型の変換クラスを対象にしておらず,同様の先行研究においては入力に対して線形な変換のみを扱っていた.今回の研究成果ではより高い表現力をもつストリーム型木トランスデューサのクラスについて透過性判定が決定可能になることを示している.この結果は(1)で示したクラスの表現力同値性を利用したものである.

この他の成果として,研究代表者のテーマの一つであった木構造データに対する双方向変換についての研究も並行して進めた.双方向変換は,二つのデータ間を同期するために必要な二つの変換のことであるが,それらの一貫性が決定可能となる変換のクラスを特徴づける計算モデルを提案した.

[令和3年度]

最終年度にあたる令和 2 年度において,形式木言語理論における木構造データ変換のモデルであるストリーム型の木トランスデューサに着目して研究を進めていたが,予期しない成果発表の延期や関連する意外な応用の発見もあったため,期間を延長して研究を進めた. 発見した応用の一つは双方向変換の一貫性の判定方法に関するものであったが,引き続き研究を進めることで判定方法の精査を行った.

具体的には,双方向変換の満たすべき性質として知られていた 12 個の規則に対し,組合せによる含意関係や同値関係を網羅的に確認し,それ以外に関係があり得ないことも併せて証明した. この事実については,さらに定理証明支援系 Coq により機械的に検査できる形で形式的に証明を行っており,その結果を強固なものとしている.

また,この成果を得る過程において,既存の双方向変換言語がもつ表現力が自明でないことを発見し,その精査も進めた.双方向変換言語はその名の通り双方向変換を実現するための言語であるが,その一貫性を自動で保証するためにさまざまな制約が課されているため,表現力が制限されている可能性がある.これを明らかにするために双方向変換のための計算モデルの設計に着手した.まだ,完全なものはできていないものの,双方向変換のもつべき一部の性質について,計算モデルの作成に成功した.本研究については次年度以降の後続プロジェクトである「双方向変換言語のための計算モデルとプログラミング言語への応用」においても,引き続き進める予定である.

5 . 主な発表論文等

〔雑誌論文〕 計15件(うち査読付論文 15件/うち国際共著 3件/うちオープンアクセス 9件)

〔雑誌論文〕 計15件(うち査読付論文 15件/うち国際共著 3件/うちオープンアクセス 9件)	
1 . 著者名	4.巻
Nakano Keisuke	215
2 . 論文標題	5.発行年
Time-symmetric Turing machines for computable involutions	2022年
3.雑誌名 Science of Computer Programming	6.最初と最後の頁
掲載論文のDOI(デジタルオブジェクト識別子)	査読の有無
10.1016/j.scico.2021.102748	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
1 . 著者名 Asano Yasuhito、Cao Yang、Hidaka Soichiro、Hu Zhenjiang、Ishihara Yasunori、Kato Hiroyuki、 Nakano Keisuke、Onizuka Makoto、Sasaki Yuya、Shimizu Toshiyuki、Takeichi Masato、Xiao Chuan、 Yoshikawa Masatoshi	4.巻 1457
2 . 論文標題	5 . 発行年
Bidirectional Collaborative Frameworks for Decentralized Data Management	2022年
3.雑誌名 Communications in Computer and Information Science book series	6.最初と最後の頁 13~51
掲載論文のDOI(デジタルオブジェクト識別子) 10.1007/978-3-030-93849-9_2	査読の有無有
オープンアクセス	国際共著
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	該当する
1.著者名	4.巻
Keisuke Nakano	46
2.論文標題	5 . 発行年
Idempotent Turing Machines	2021年
3.雑誌名	6.最初と最後の頁
International Symposium on Mathematical Foundations of Computer Science	79:1~79:18
掲載論文のDOI(デジタルオブジェクト識別子) 10.4230/LIPIcs.MFCS.2021.79	 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
1 . 著者名	4.巻
Nakano Keisuke	13
2 . 論文標題	5.発行年
A Tangled Web of 12 Lens Laws	2021年
3.雑誌名	6.最初と最後の頁
Reversible Computation	185~203
掲載論文のDOI(デジタルオブジェクト識別子)	査読の有無
10.1007/978-3-030-79837-6_11	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

1 . 著者名 Ishihara Yasunori、Kato Hiroyuki、Nakano Keisuke、Onizuka Makoto、Sasaki Yuya	4 . 巻
	- - - 25/2/F
2 . 論文標題 Toward BX-Based Architecture for Controlling and Sharing Distributed Data	5 . 発行年 2019年
3.雑誌名 BigComp 2019	6.最初と最後の頁 1-5
掲載論文のDOI(デジタルオブジェクト識別子) 10.1109/BIGCOMP.2019.8679145	 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 Asano Yasuhito、Herr Dennis-Florian、Ishihara Yasunori、Kato Hiroyuki、Nakano Keisuke、Onizuka Makoto、Sasaki Yuya	4.巻
2 . 論文標題 Flexible Framework for Data Integration and Update Propagation: System Aspect	5.発行年 2019年
3 . 雑誌名 BigComp 2019	6.最初と最後の頁 1-5
掲載論文のDOI(デジタルオブジェクト識別子) 10.1109/BIGCOMP.2019.8679236	 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 該当する
1.著者名 Keisuke Nakano	4.巻
2 . 論文標題 Towards a Complete Picture of Lens Laws	5 . 発行年 2019年
3.雑誌名 Proceedings of the Third Workshop on Software Foundations for Data Interoperability	6.最初と最後の頁 1-6
 	 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
1 . 著者名 Mirai Ikebuchi, Keisuke Nakano	4.巻
2 . 論文標題 On Repetitive Right Application of B-Terms	5 . 発行年 2018年
3 . 雑誌名 International Conference on Formal Structures for Computation and Deduction (FSCD 2018)	6.最初と最後の頁 18:1-18:15
引載論文のDOI(デジタルオブジェクト識別子) 10.4230/LIPIcs.FSCD.2018.18	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

1 . 著者名	4.巻
高橋 祐多、中野 圭介	35
2 . 論文標題	5 . 発行年
木から文字列への決定性トップダウン変換の等価性判定アルゴリズムの実用性について	2018年
3.雑誌名	6.最初と最後の頁
コンピュータ ソフトウェア	52-71
掲載論文のDOI (デジタルオブジェクト識別子) 10.11309/jssst.35.52	 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
1 . 著者名	4.巻
Kazuhiro Abe, Keisuke Nakano	25
2.論文標題	5 . 発行年
Towards Practical Typechecking for Macro Forest Transducers	2017年
3.雑誌名	6.最初と最後の頁
Journal of Information Processing	962-974
掲載論文のDOI (デジタルオプジェクト識別子)	査読の有無
10.2197/ipsjjip.25.962	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
1 . 著者名	4.巻
Daisuke Kinoshita, Keisuke Nakano	1827
2.論文標題	5 . 発行年
Bidirectional Certified Programming	2017年
3.雑誌名	6.最初と最後の頁
CEUR Workshop Proceedings	31-38
掲載論文のDOI(デジタルオブジェクト識別子) なし	 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15件(うち招待講演 1件/うち国際学会 8件) 1.発表者名 Keisuke Nakano	
2 . 発表標題 On Turing machines with syntactic restrictions	

School of Computer and Cyber Sciences (CCS) Colloquium Series, Augusta University (招待講演) (国際学会)

3 . 学会等名

4 . 発表年 2021年

1. 発表者名
Keisuke Nakano
2 . 発表標題
Idempotent Turing Machines
3.学会等名
3 . 子云寺ロ 46th International Symposium on Mathematical Foundations of Computer Science(国際学会)
40th International Symposium on Mathematical Foundations of Computer Science (国际子云)
4.発表年
2021年
20217
1.発表者名
Keisuke Nakano
Nersune Manario
2.発表標題
A Tangled Web of 12 Lens Laws
3.学会等名
13th Conference on Reversible Computation (RC 2021)(国際学会)
4.発表年
2021年
1.発表者名
佐藤 光
2 PV == 1 TR PR
2.発表標題
マクロ木変換器を融合するための高階木変換器の低階化
3.学会等名
コ・デムサロ 日本ソフトウェア科学会 第38回大会
ロボンノー フェブ 11 子 ム おの日 八 ム
4 . 発表年
2021年
1.発表者名
Keisuke Nakano
2. 発表標題
Involutory Turing Machines
3.学会等名
12th Conference on Reversible Computation (RC 2020)(国際学会)
4. 発表年
2020年

1.発表者名
Yuta Takahashi
2.発表標題
Streaming Ranked-Tree-to-String Transducers
a WAR
3.学会等名
Conference on Implementation and Application of Automata(国際学会)
4.発表年
2019年
1.発表者名
中野圭介
2.発表標題
モデル検査のパズル化
3.学会等名
夏のプログラミング・シンポジウム 2019
4.発表年
2019年
1.発表者名
Keisuke Nakano
2 . 発表標題
On Repetitive Right Application of B-Terms
3.学会等名
International Conference on Formal Structures for Computation and Deduction (FSCD 2018)(国際学会)
4.発表年
2018年
1.発表者名
Mirai Ikebuchi
2 . 発表標題
ComplCoq: Rewrite Hint Construction with Completion Procedures
3 . 学会等名
The Coq Workshop 2018 (国際学会)
4.発表年
2018年

1.発表者名
中野 圭介
2 . 発表標題
On Repetitive Right Application of B-Terms
3.学会等名
3 . 子云寺石 第21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
カ41回ノロノノへングのよびブログブミング音画フ ^ー ブグコップ
4.発表年
- 2019年
±010 ⁻ T
1.発表者名
・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・
1001回)注入
2.発表標題
数値や文字列に対するガードを含む累積引数付き再帰関数の融合
3. 学会等名
第21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
2019年
1.発表者名
高橋 祐多
2. 発表標題
ランク付き木から文字列への決定性ストリーム変換器の表現力
2
3. 学会等名
第21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
2019年
1. 発表者名
Keisuke Nakano
2.発表標題
2 . 光花標題 On the repetitive right application of B-terms
on the reporterior right appropriation of beteine
3.学会等名
Dagstuhl Seminar on Formal Methods of Transformations(国際学会)
C
4 . 発表年
2017年

1.発表者名 山田 伊織,中野 圭介	
2.発表標題 Coqにおける手続き的証明から宣言的証明への変換	
3.学会等名 第20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 (PPL2018)	
4 . 発表年 2018年	
1.発表者名 西山舜,中野圭介	
2 . 発表標題 スタック構造の累積引数を持つ関数を融合するための木変換器合成	
3.学会等名 第20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(PPL2018)	
4. 発表年 2018年	
〔図書〕 計1件	
1.著者名 Daniel P. Friedman, Carl Eastlund, 中野圭介	4 . 発行年 2017年
2.出版社 ラムダノート社	5.総ページ数 240
3.書名 定理証明手習い	
〔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) (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	

相手方研究機関

共同研究相手国

米国	Massachusetts Institute of Technology		
中国	Peking University		