研 究 終 了 報 告 書

「階層的グラフの書き換え系での文脈等価性証明支援」

研究期間: 2019 年 10 月～2022 年 3 月
研 究 者: 室屋 晃子

1. 研究のねらい

コンピュータプログラムは、プログラミング言語が提供する様々な機能を組み合わせることによって作り上げられる。それぞれの機能には期待される振る舞いがあるが、機能同士の組み合わせ次第では期待通りの振る舞いが得られない状況が発生することもある。

例えば、関数型プログラミング言語における最も基本的な振る舞いとして、図 1 の 2 つのプログラム片 M, N は同じ実行結果を与えると期待される。実際にこれらをプログラム文脈 C において実行すると 2 つのプログラム C[M], C[N]の実行結果は等しくなる。しかしこれらを、メモリ使用状況を取得する機能（stat 関数）と組み合わせると、期待される振る舞いは得られない。実際に 2 つのプログラム C'[M], C'[N]の実行結果は異なってしまう。

プログラム片の例

```
M = (fun x -> x + x) 2
N = 2 + 2
```

M,Nを区別しないプログラム文脈

```
C = ([]) * 3
```

M,Nを区別するプログラム文脈

```
C' = stat([])
```

図 1

本研究では、このように期待される振る舞いが必ずしも得られない状況に対する数理的理解を助けることを目的として、様々な言語機能同士を組み合わせた時にそれらが期待通りに振る舞うことを保証する数理的手法を開発する。具体的には、言語機能の振る舞いを等式として定式化する「文脈等価性」という概念に対して、言語機能に依らない一般的な証明手法を構築し、さらにその部分的な自動化を目指す。

文脈等価性の証明によって保証できることに、例えばコンパイラ最適化の正しさがある。つまり、プログラム実行の高速化を目的としてコンパイラが行うプログラムの部分的変更が、プログラム全体の振る舞いを変化させないことを、その部分的変更の文脈等価性を示すことにより保証することができる。文脈等価性の汎用的かつ自動化可能な証明手法は、あるプログラミング言語のコンパイラ最適化の正しさを保証するだけでなく、例えばそのプログラミング言語の機能を追加・変更した際にも、既存のコンパイラ最適化が引き続き正しいことを簡便に確かめるためにも用いることができる。本研究はそのような証明手法の構築を目指すことにより、コンパイラ開発、特にコンパイラ最適化の安全性保証に対する貢献をも狙うものである。

2. 研究成果

(1)概要

文脈等価性の汎用的かつ自動化可能な手法を目的として、既に得られているプロトタイプ的な証明手法（以下では「*ベース手法*」と呼ぶことにする）の問題点を整理し、それらを改善することを目標に研究を遂行した。

このベース手法は、階層的グラフの書き換え系としてプログラムの実行をモデリングすることで、共通部分グラフの存在に着目することで示せるような文脈等価性の十分条件を与えるものである。しかしこの手法には、十分条件を示す方法が未整備であり特に自動化の方針が明らかでないこと、プログラムの決定的な振る舞いしか対象でなく汎用性がないこと、さらにこれら2つの問題点を解決するには数学的基盤が弱いこと、といった問題点がある。そこでまず数学的基盤を整備し、それから汎用化や自動化を追求する方針で研究を始めた。

数学的基盤の整備に当たり圏論を用いた既知のグラフ書き換え系の定式化を活用することを検討したが、非常に単純化された状況から理論を積み重ねる必要があり、当初予想よりも時間が必要であることが分かった。そこで階層的グラフ書き換え系によるモデリングの理論的整備を待たずに、ベース手法の技術的な核である文脈等価性の十分条件をより数学的に整ったモデルに対して定式化し追究するという方針に変更した。

具体的には、状態遷移系理論からのアプローチで十分条件を「模倣関係」の一種として定式化することに取り組み、証明手法の汎用化や自動化を目指した。この定式化を非決定的オートマトンに対して与えることで非決定的なプログラムの振る舞いを扱えるようにしたほか、得られた模倣関係が計算可能になりうることも示し、文脈等価性の実際の証明手法を構築するには至らなかったものの、一定の進展を得ることができた。

また、項書き換え理論からのアプローチで十分条件を定式化する取り組みも行った。上記の状態遷移系理論からのアプローチに注力したこともあり具体的な成果を得るには至らなかったが、十分条件のアイデアが、項書き換え理論での基礎的な問題である合流性問題や停止性問題との類似性を持つという重要な観察を得た。

これらの取り組みでは、文脈等価性の汎用的かつ自動化可能な証明手法を実際に構築し研究目的を達成するには至らなかった。しかし技術的な核となる模倣関係を、当初想定していた階層的グラフの書き換え系というモデルより抽象的かつ一般的なモデルに対して理論的に整備した。さらに、その具体的成果が、ベース手法を目的とする汎用的かつ自動化可能な手法に近づけるために有効であることが確認できた。

（2）詳細

プログラムの実行結果が一意に定まるような言語機能に対して適用可能な、文脈等価性のプロトタイプ的な証明手法（以下では「*ベース手法*」と呼ぶことにする）が既に得られており、これを改良および拡張することを目標として研究を行った。このベース手法の特徴は、プログラムの実行過程を階層的グラフの書き換え系として図2のようにモデリングすることである。グラフ表現に基づくこのモデルでは、文脈等価性の一般的な証明が共通部分グラフの存在に着目することで可能になる。

	グラフ書き換えモデルでの表現
プログラム	階層的グラフ
プログラム実行のステップ	部分グラフの書き換え
各言語機能の振る舞い	部分グラフの書き換え規則

図2

具体的にベース手法は、2 つのプログラム片の表現である 2 つの階層的グラフの組 *(G, H)*

が文脈等価性を満たすことを、以下のような十分条件を示すことに帰着させる。

> 十分条件★ 差異が *(G,H)* で与えられるような任意の階層的グラフの組 *(P, Q)* について、*P* の書き換えの数ステップは *Q* の書き換えの数ステップで模倣できる（かつ、同様に *Q* の書き換えの数ステップも *P* の書き換えの数ステップで模倣できる）

そして共通部分グラフの存在に着目することで、この十分条件が示せるというものである。

本研究期間においては、まずベース手法の問題点を以下のように整理した。

1. 十分条件★を示す方法が未整備であり、煩雑な証明を手作業で行う必要がある。
2. 扱えるプログラムの振る舞いに、決定的であるという制限があり、例えば非決定的あるいは確率的な振る舞いが扱えない。
3. 上の問題点 1 および問題点 2 を解決するには、階層的グラフ書き換え系によるプログラム実行のモデリングの数学的基盤が弱い。ベース手法の構築に当たっては、状態遷移系・書き換え理論・グラフ理論などの複数の分野から着想を得た様々な技法が複合的に用いられており、中には新奇と思われる手法も混じっている。

そして、初めに問題点 3 に取り組み、ベース手法で使用するプログラム実行のモデルの数学的基盤を整備してから、問題点 1 の解決を通じてベース手法の自動化を図るとともに、問題点 2 を解決して適用可能なプログラムの振る舞いを拡大する、という方針で研究を始めた。しかしモデルの数学的基盤整備には技術的要素が多重に要求され、非常に単純化された状況から理論を積み重ねる必要があり、当初予想よりも時間が必要であることが分かった。

そこで階層的グラフの書き換え系というモデルの理論的整備を待たずに、ベース手法の技術的な核である十分条件★をより数学的に整ったモデルに対して定式化し追究するという方針に変更した。具体的には、状態遷移系理論や項書き換え理論からのアプローチで問題点 1 および問題点 2 を解決することに注力した。

本研究期間中の具体的な取り組みと成果は以下の通りである。

(A) ベース手法が使用するモデルの数学的基盤の整備 – 問題点 3 への取り組み

ここでは主に 2 つの方向性の取り組みを行った。第一に、ベース手法およびその具体的な使用例を論文にまとめる作業を行った。ベース手法自体は当該研究者の博士論文（University of Birmingham, 2020）の内容に基づくが、ここでは新たにベース手法の使用フローのうち具体例に依らない部分と一般性のある基盤的な部分の切り分けを行った。第二に、グラフ書き換え系の既知の定式化を用いたベース手法の再構築を模索した。特に圏論による定式化を検討したが、プログラムの構文的複雑さに応じて圏論での構成も複雑になること、そして非常に単純なプログラムに対しても圏論的定式化は容易でないという観察を得たため、階層的グラフの書き換え系というモデルに拘らずに問題点 1 および問題点 2 に注力するという研究方針の変更に至った。また自動化可能な定式化に関する知見を得るため、グラフ書き換え系を実装するためのソフトウェアを開発している上田和紀教授（早稲田大学）を訪問しセミナー発表や意見交換を行った。

(B) 状態遷移系理論による十分条件★の定式化 – 問題点 1 および問題点 2 への取り組み

十分条件★は、状態遷移系の比較に用いられる「模倣関係」の変種として定式化することができる。問題点 2(ベース手法の適用範囲の制限)は、ベース手法が用いるこの特殊な模倣関係がプログラムの決定的な振る舞いに特化していることによる。そのため、この模倣関係をより広範な振る舞いを扱えるように一般化する研究を行った。具体的には、ベース手法が用いる階層的グラフ書き換え系を抽象化および一般化したとみなせる非決定的オートマトンに対して、ベース手法が用いたものと同様の性質を持つ模倣関係を与えた。得られた模倣関係(以下では「*前順序つき模倣関係*」と呼ぶ)は、オートマトンの振る舞いの比較方式を受理語上の前順序で拡張していることが新奇なものである。

この前順序つき模倣関係は、問題点 2 の解決という目標に向けて不完全ながら一定の進展を与えるものである。つまり、ベース手法が扱えなかった多様なプログラムの振る舞いのうち、非決定的な振る舞いの扱いが可能になった。一方で、確率的な振る舞いを扱うには不十分であることも明らかになった。

また、前順序つき模倣関係は、前順序が計算可能かつ非決定的オートマトンが有限状態ならば、計算可能であることを示した。前順序の有用な例として調査したものは全て計算可能であったことから、これはプログラム実行のモデルが有限状態ならば、十分条件★が成り立つか否かを計算可能であることを意味する。具体的な計算アルゴリズムを与えるには至らなかったものの、この結果は問題点 1 の理論的な解決につながるものである。

この取り組みは模倣関係の専門家である卜部夏木氏(NII)と共同で行った。研究成果のコア部分(つまり前順序つき模倣関係の定義、およびその定義の正当性の証明)には定理証明支援系 Agda での実装という形で裏付けを与え、その際に所属研究グループの博士後期課程学生を研究補助者として雇用した。得られた結果を論文にまとめるとともに、部分的な成果は国際ワークショップ CMCS 2020 での招待講演[その他の成果 1]や国際会議 CALCO 2021 の Early Ideas トラックで発表した[その他の成果 2]。

(C) 項書き換え系理論による十分条件★の証明 – 問題点 1 への取り組み

ベース手法が用いる模倣関係を実際に構成する過程では、階層的グラフの書き換え系というモデルの性質上、注目する部分グラフに対してどのような書き換え(変化)が可能であるかを解析することが不可欠である。この解析は、注目する部分グラフと、書き換え対象になり得る部分グラフとの可能な重なり方を全列挙することでなされる。項書き換え系理論の専門家である浜名誠准教授(群馬大学)との議論を通じて、このような部分グラフ同士の重なり方の列挙に必要なアイデアは、項書き換え理論での基礎的な問題である合流性問題や停止性問題との類似性を持つという観察を得た。研究の全体方針の変更以降は主に上記 (B) に注力したこともありこの取り組みで具体的な成果を得るには至らなかったものの、この観察を元に、特殊な合流性問題を解くことで模倣関係を構成するような手法の構築を模索した。

上記の 3 つの取り組みでは、文脈等価性の汎用的かつ自動化可能な証明手法を実際に構築

し研究目的を達成するには至らなかった。しかし技術的な核となる模倣関係を、当初想定していた階層的グラフの書き換え系というモデルより抽象的かつ一般的なモデルに対して理論的に整備した。さらに、その具体的成果が、ベース手法を目的とする汎用的かつ自動化可能な手法に近づけるために有効であることが確認できた。

3.　今後の展開

今後は取り組み（B）と取り組み（C）を継続し、本研究の目的の達成を引き続き目指す。

第一に、取り組み（B）で得られた成果（つまり前順序つき模倣関係）を理論的に発展させることで問題点 2 の解決を目指す。また、前順序つき模倣関係を用いてベース手法を再構築することで、プログラムの多様な振る舞いを扱える文脈等価性の証明手法を与えることを目指す。ここでは上記の研究の結果を用いることで証明手法の一般性を達成するだけでなく、模倣関係の計算アルゴリズムを与え実装することにより証明手法を自動化することを目指す。

第二に、本研究期間中には具体的成果に至らなかった取り組み（C）を推進し、これを通じて文脈等価性の一般的かつ自動化可能な証明手法を与えることを目指す。

4.　自己評価

研究目的の達成状況について

取り組み（B）により技術的な核となる模倣関係を定式化することで、問題点 1 および問題点 2 の双方の解決に向けた一定の進展が得られた。これにより、目的とする文脈等価性の証明手法の実際の構築には至らなかったが、その技術的な核となる部分の理論的整備という成果を得られた。目的の達成に至らなかった要因は主に時間的であると捉えている。つまり、取り組み（A）での観察から研究方針を変更するまでに研究期間の半分ほどを要したことで、特に取り組み（B）により目的を達成するための十分な時間が得られなかったと考える。

研究方針の変更までの時間は、洗い出した 3 つの問題点から具体的な研究テーマ（取り組み（B）と取り組み（C））を抽出するのに要した時間でもあった。ここでは研究テーマ抽出の困難さに直面し、研究期間の小さくない部分を費やしてしまったが、自立した研究者としてのキャリアが浅い段階においては必要かつ有意義な挑戦であったと考えている。

具体的な研究テーマの抽出を通じて、階層的グラフの書き換え系モデルに基づくベース手法のアイデアが、より抽象的なモデルであるオートマトンや、プログラム意味論においてはより典型的な項書き換え系モデルにも適用可能であることを明らかにする方向性に研究が進むこととなった。これは研究目的を外れるものではなく、むしろ得られる手法をより強力にすることが期待される。この点では、階層的グラフの書き換え系に焦点を当てていた当初の想定より発展的な形での研究成果につながったといえる。

研究の進め方について

研究を進めるにあたり、特に取り組み（B）と取り組み（C）においては国内の専門家との議論や協働が大きな推進力になった。特に研究テーマ抽出に注力した期間は、広く研究動向調査や意見交換を行うために学会等に参加することが有用と考え、そのために研究費を活用する予定であったが、パンデミックの影響としてこのような活動は制限を受ける形となった。

また取り組み（B）の進展により、当初予定にはなかった研究補助者の雇用を行った。研究成果を定理証明支援系で裏付けすることは有用であり、研究補助者の雇用によりこれを大きく促進することができた。また研究補助者としての博士後期学生と協働するという経験は、指導学生を持たない立場にある現状では特に有意義であった。

研究成果の波及効果について

本研究では文脈等価性の証明というプログラム意味論の問題に取り組む過程で、状態遷移系理論や項書き換え系理論といった関連分野にそれぞれ貢献しうる成果や知見を得ることができた。まず、取り組み（B）で得られた模倣関係は文脈等価性の証明を目的に与えたものであるが、模倣関係の典型的な応用先であるモデル検査に対して、新奇な応用の可能性を示唆する観察を得ている。また取り組み（C）は具体的成果に至らなかったものの、項書き換え系理論とプログラム意味論との相互作用を推進する効果が見込まれる。これら二つの理論分野は、書き換えを静的性質と捉えるか動的性質と捉えるかといった差異などから、知見や技術の移植が進んでいないという現状がある。取り組み（C）は、項書き換え系理論の技術をプログラム意味論に応用する実例を与えることで、両分野の知見を融合・発展させる研究を加速できると期待される。

5.　主な研究成果リスト

（1）代表的な論文（原著論文）発表

研究期間累積件数：0件

（2）特許出願

研究期間全出願件数：0 件（特許公開前のものも含む）

（3）その他の成果（主要な学会発表、受賞、著作物、プレスリリース等）

1. 国際ワークショップ CMCS 2020 での招待公演（2020 年 10 月 12 日）
   Koko Muroya. “Hypernet Semantics and Robust Observational Equivalence.”
2. 国際学会 CALCO 2021 の Early Ideas トラックでの査読付き口頭発表（2021 年 9 月 2 日）
   Koko Muroya, Takahiro Sanada and Natsuki Urabe. “Preorder-Constrained Simulation (Early Ideas).”