

A Framework for Construction of Correct Parallel Programs based on Distributed Specialization Generation

(特殊化生成の分散処理に基づく正当な並列プログラム構築のための枠組)

学位論文内容の要旨

本研究の目的は正当な並列処理プログラムの自動生成の基盤となるようなプログラム構築のための枠組を提案することである。この枠組ではプログラム構築を以下の三ステップから行う: 最初に確定節集合により問題を宣言的に定義する。次に計算を確定節の書換え (変換) と見なし、適用される正しい変換をルールや関数として表現する。正しい変換とは確定節集合の意味を保存する変換 (等価変換) である。そして最後にルールなどにより定義された正しい変換結果をもたらず並列処理プログラムを構築する。

この枠組は、逐次処理プログラムを対象とし同じ目標を目指す研究である等価変換理論に基づくプログラム構築の枠組 (ETF) を拡張したものである。ETF では問題定義からプログラムを直接生成するのではなく、正しい計算を定義するルールが導入される。そしてルールの生成・改善・検証をルール単位で行い、それらのルールを組み合わせることで、多様なアルゴリズムに基づく正当なプログラムの構築を可能とする。ただし ETF の正当性に関する理論は逐次処理 (ルールの逐次適用) を前提としているために直接並列処理への拡張はできない。並列処理の場合に、非同期に動作するプロセス上の各プログラムが全体として正しい計算結果をもたらずことを保証するのは困難である。本研究では宣言的記述による問題定義、等価性に基づく正当性の定義、ルールによるモジュラリティを持った計算定義を基礎にこれまでの ETF を発展させたことによって、並列計算機上で動作する正当なプログラム構築のための理論と枠組を提案した。

拡張はルール集合による正しい計算の定義とそれを利用したアルゴリズムの実装であるプログラムに対して行われる。今回提案する枠組では等価変換ルールに基づく計算において重要な役割を果たす「特殊化」に関する計算を分散処理する枠組「特殊化の分散生成による並列処理 (DSG-PP)」を採用した。特殊化は論理プログラミングの代入を拡張した概念であり、ETF では「特殊化システム」として拡張可能である。DSG-PP は変数に束縛されるべき値の計算の並列処理であり、束縛されるべき正しい値を定義する関数とルールを計算定義に追加できるように ETF を拡張した。

これらの関数やルールは、逐次的に適用した場合には正当性の条件を満たすように定義されている。しかしその計算の一部を分散して処理した場合の正当性は保証されない。それに対して本論文では、並列に処理された結果を矛盾なく利用するためのアルゴリズムを組み込むように実装プログラムを拡張した。分散処理はマスタ・ワーカモデルを採用し、計算結果の正当性はマスタが管理する任意の時点の確定節集合が、問題定義と等価であることを証明することで保証した。

既存の正当なプログラムを構築する方法として、生成したプログラムの正当性を問題定義に照らして検証するアプローチが考えられる。しかしこのアプローチによる自動生成では、検証すべきプログラムをいかに生成するかが問題となる。並列処理プログラミングでは既存の逐次処理アルゴリズムを選択し、そこに含まれる並列性に基づいてプログラムを生成することが多い。この場合、並列計算機のための新たなアルゴリズムに基づくプログラムの生成の可能性を失う。ランダムに行った場合は正当なプログラムを得るまでの無駄が多くなる。利用可能な計算資源などに対して最適なプログラムや、新たなアルゴリズムに基づくプログラムを抽出するためには、多くの正当なプログラムを効率的に生成する必要があるがこれらのアプローチでは困難である。それに対して本論文で提案

した枠組は最初から正当なプログラムを生成するアプローチであり、新たなアルゴリズムに基づく並列処理プログラムの生成にとって有利な枠組である。

本論文の構成は以下の通りである。

第1章は本論文の目的と背景を解説する。第2章は並列処理プログラム生成の概説、および本研究でターゲットとする並列計算アーキテクチャを導入する。第3章ではDSG-PPの正当性の理論的基礎となる特殊化システム、等価変換、そしてETFについて説明する。ETFでは問題は宣言的に、つまり問題と正しい解の関係として定義される。等価変換理論における計算、つまり質問を示す確定節(query)の変換は特殊化と節内のボディアトムの変換から構成される。ここでその特殊化を定義する特殊化システムを導入する。

第4章ではこの特殊化に基づく並列処理の概説を行う。DSG-PPではマスタがqueryの変換を担当する。ワーカはマスタからqueryのボディアトムの一部を受け取り、queryへの正当な特殊化に関する情報(特殊化情報)を計算する。マスタはワーカから受け取った特殊化情報をqueryに適用する。このDSG-PPの結果の正当性にとって問題となるのがマスタからワーカへのボディアトム送信とその結果である特殊化情報受信の時間差であることが示される。この時間差の間にマスタの計算状態が変化する可能性があり、受信した特殊化情報による特殊化の結果の正しさは自明ではないからである。

第5章では正当な特殊化のために、受信した特殊化情報をマスタで行われた変換の履歴(Memo Tree, MT)により「補正」するアルゴリズムを組み込んだDSG-PPを提示する。この補正に基づく正当な並列処理のためのアルゴリズムをDSG-PP(MT)と呼ぶ。しかしMTは処理が進むと共に単調に増加することが問題となる。

解決方法として第6章ではMTを必要としないアルゴリズム(DSG-PP(LW))を提案する。DSG-PP(LW)では特殊化情報を変数からパターンへの等価制約に限定する。等価制約とは二つの項が等しいことを示す制約であり、二つの項を等価とする特殊化の適用によって解消される。マスタは送信されたボディアトムに含まれる変数を記録する。ワーカは特殊化情報を計算し、その特殊化情報に基づいて受信したボディアトムに含まれる変数を特殊化した結果をマスタに送信する。この間にマスタ内で適用された特殊化はマスタ内に記録された変数に反映される。ワーカから受信した結果とマスタ内の記録された結果は同じ変数を特殊化した結果であり、等価でなければならない。マスタはその等価性を制約として解消することで正当な特殊化を行う。この方式では送信した変数の記録がマスタに求められるが、等価制約を解消した時点でそれは削除されるので記憶負荷は限定される。

第7章では並列処理に必要なDSG以外のコンポーネントについて説明する。実装されるプログラムにおいてはリクエストのタイミング、ワーカの管理、マスタ・ワーカ間通信の実装なども、パフォーマンスにとって重要な要素となる。

第8章ではDSG-PP(LW)に基づくプログラム構築例を通してその有効性を示す。制約充足問題を取り上げる。問題を自然に定義するための特殊化システムが導入され、次に並列処理の対象となるアトムを特殊化するためのルールや等価制約を解決するルールを含むルール集合が定義される。そしてそれらのルールと通信を管理するアルゴリズムを実行するマスタとワーカのプログラムが実装される。これによりDSG-PP(LW)の有効性を示す。

第9章では本論文で得られた知見と本研究の今後の方向性について検討する。

学位論文審査の要旨

主査	教授	赤間	清
副査	教授	栗原	正仁
副査	教授	水田	正弘
副査	准教授	棟朝	雅晴

学位論文題名

A Framework for Construction of Correct Parallel Programs based on Distributed Specialization Generation

(特殊化生成の分散処理に基づく正当な並列プログラム構築のための枠組)

与えられた問題を解くためのプログラムを正しく構築するのは、非常に難しい。それは問題とそれを正しく高速に解くプログラムの関係が科学的に十分解明されていないからである。良いプログラムを構築するには、適切なアルゴリズムを探索することが重要になるが、アルゴリズム探索を含む本格的なプログラムの自動生成の研究は、きわめて難易度の高い課題であると考えられる。

本研究では、並列計算環境のもとで動作するプログラムを自動生成するための基礎となる方法を提案している。最大の課題は、並列性をいかに引き出すかである。OR 並列の場合、独立な部分が陽に見えており、解決は容易である。AND 並列の場合でも、独立な部分が豊富にある単純な場合には、自明な並列性を持ち、並列性を引き出すことは容易である。現在利用されている並列プログラムの多くはこのような自明な並列性だけを利用しており、その範囲では正当性の確保に困難は生じない。しかし大多数の一般の AND 並列では、自明な並列性は陽には見えず、AND 並列部分の計算の絡み合いを克服して正しく効率的な並列プログラムを構築するのは容易ではない。AND 並列において十分に一般性のある並列性抽出方法を確立することは、並列プログラム構築における大きな課題であり続けている。本研究は、その重要課題に挑戦し、「特殊化情報の分散計算」という新しいコンセプトを導入して、AND 並列に対する一般性の高い並列性抽出方法を提案している。

本論文の構成は以下の通りである。

第1章で本論文の背景と目的を述べた後、第2章では、並列処理プログラムを生成する枠組みや本論文で採用する並列計算アーキテクチャ(マスタ・ワーカ型)について記述している。

第3章では対象とする問題を規定するための諸概念、すなわち、特殊化システム、アトム、節、宣言的記述、宣言的記述の意味や、考察する問題の数学的定義、得られるべき解答などを導入している。また、計算アルゴリズム構築のための重要な構成要素である等価変換ルールの定義とその表記方法について議論している。

第4章では、特殊化の分散計算に基づく並列計算の主要なコンポーネントを導入している。質問節を変換するのはマスタであり、任意個のワーカはそのための情報を作り出す計算を並行して行う。マスタは質問節の情報として、質問節の要素であるアトムをいくつかワーカに送る。ワーカは受け取ったアトム集合に基づいて計算を行い、特殊化情報を得る。それはマスタに送られ、マスタによって質問節に適用され、マスタの計算は促進される。

この章では、ワーカの計算が満たすべき条件も与えている。これは十分緩い条件であり、生成できる並列プログラムの多様性を確保するのに役立つ。ワーカの特 specialization 情報をマスタが適用する部分に対する条件なども与えている。しかしこれらの条件が満たされても、並列計算の正当性は必ずしも確保できない。マスタやワーカなどの多数の計算が絡み合うことがその原因であることが説明されている。

第5章では、並列プログラムを構成する第1の方法を導入している。その特徴はマスタが specialization 情報のツリー状の履歴を保持し更新しながら利用することである。マスタが受信した specialization 情報は、この履歴をもとに補正され、質問節に適用される。

第6章では、並列プログラムを構成する第2の方法を導入している。これはマスタがツリー状の履歴を持たなくてもよい方法であり、マスタの空間的ならびに時間的コストを軽減することができる。

第7章では、提案方法に基づいて並列プログラムを構築する方法の概要を示している。並列プログラムの最終的なパフォーマンスを向上させるために、ワーカへのリクエストのタイミングや利用できるワーカの管理、マスタ・ワーカ間通信の実装なども記述されている。

第8章では提案方法を制約充足問題に適用した例を示し、並列計算の効果を評価するための実験を行っている。

第9章では、本論文で得られた結果をまとめている。

これを要するに、本論文は AND 並列からの一般性の高い並列性抽出方法を考案し、宣言的記述で書かれた広範な問題クラスに含まれる問題が与えられたとき、それを解く正当で高速な並列プログラムを多様に自動生成するための基礎を提案したものであり、人工知能およびソフトウェア科学の分野に貢献するところ大なるものがある。よって著者は、博士(情報科学)を授与される資格あるものと認める。