

論文の内容の要旨

論文題目「並列処理時間のモデル化による性能評価方法に関する研究」

学位申請者 折居 茂夫

キーワード：並列処理，性能評価，並列効率，性能予測，モデル化

電子計算機による高速計算の歴史は、1940年代に開発されたABCとENIACから始まり70年を超える。この間に素子は、真空管、トランジスタ、集積回路と変遷し、更に半導体テクノロジーの向上により計算速度の向上が図られてきた。一方、個々の原子の動きが見えるほど回路の集積度が高まった現在、集積度の向上で計算速度を向上することに限界が来つつある。この限界を超える方法として、独立に計算できる部分を同時に実行して計算速度の向上を図る並列処理がある。並列計算機の歴史は古く、1960年代には最初の並列計算機と言われるILLIACIVが開発されている。一方高速計算の歴史の中で並列計算機が主流になったのはここ20年に満たない。このように利用が遅れた大きな理由の一つは、並列計算機が処理の「独立に計算できる部分」を利用して高速化を図ることにある。これは処理毎に性能が異なることを意味し、また実際に期待した性能を得ることが容易でなかった。従って半導体テクノロジーの進歩で計算機の高速化が達成できた時代には、並列計算機は積極的に利用されなかった。一方半導体テクノロジーの進歩の限界が予見されるようになった1990年代の中ごろ以降の高速計算では、並列処理が積極的に取り入れられるようになった。現在世界トップレベルの性能を持つ「京」スーパーコンピュータも並列計算機である。

並列計算機では多くのプロセッサを用いることにより大規模な問題の計算が可能となる。一方他のプロセッサにあるデータをネットワーク経由で授受するため、このデータアクセス時間がオーバーヘッドとなり、高速計算を妨げるボトルネックとなる。このオーバーヘッドは通信方法等により変化し、また同一プログラムであっても計算する問題の大きさにより変化する。このように並列処理を用いて高速計算を追及した結果、高速計算の性能は、並列計算機とプログラムの組み合わせ、更にプログラムが実行する問題の規模等により著しく異なることになった。

このような状況下では性能の評価が重要である。ところが従来から用いられてきた、大多数の並列処理に使える汎用的な性能評価指標である並列効率の決定には、1プロセッサの処理時間を基準とするため、同じ計算を2度行う必要があった。従ってこの性能評価指標は日常の並列処理を評価する指標に使えないという欠点があった。そこで汎用性を保ちつつこの問題を解決する新しい効率の性能指標を提案した。この提案により日常の並列処理の性能を評価できるようになった。

並列処理ではまた、プロセッサ数や問題の規模を大きくした場合の性能予測が重要である。処理時間のモデル化はこの課題に対する主な解決方法の一つである。そこで処理時間の殆どを占めるプログラム中のカーネルの処理時間をモデル化し、プログラムの処理時間

をモデル化する方法を提案した。この方法はいわゆるホワイトボックスアプローチで、プログラムの構造に基づいた予測ができる利点を持つ一方、モデルパラメータが多くその値の決定の容易さに課題を持つ。そこでプログラム実行時の処理時間を多項式で現わすブラックボックスアプローチでモデル化する方法を提案した。このモデル化方法は、複雑なプログラムでも簡単にモデル化できる利点を持つが、予測精度を上げるにはモデルパラメータ決定に多くの観測データを必要とした。そこで与えられた基底関数モデルから非負モデルパラメータを推定するアルゴリズムを提案した。このアルゴリズムにより、モデル化で生じる過剰適合を抑制してモデルの予測精度の向上を図ることができるようになった。

本学位論文の構成を以下に示す。

第1章では本研究の背景として、高速計算が現在の並列計算に至るまでを概説し、並列処理の性能評価についての本研究の提案に至った経緯について説明を与えている。

第2章では処理時間モデルを用いて並列処理の性能に関して説明した。はじめに処理時間モデルを分類し、第一に性能評価指標を導出した。第二にホワイトボックスアプローチとブラックボックスアプローチから処理時間をモデル化する方法を提案した。第三にモデルパラメータに非負制約を付けて推定することにより、処理時間モデルの予測精度を向上するアルゴリズムを提案した。

第3章では、第2章で提案した性能評価指標を用いて実際の並列処理の性能の評価を行った。提案した並列効率の性能評価指標とサブ性能評価指標を用いて、分散メモリ型並列計算機で行った並列処理と共有型並列計算機で行った並列処理の性能評価ができることを確認した。また分散メモリ型並列計算機のハードウェアの効率評価ができることを確認した。

第4章では、第一にホワイトボックスアプローチで並列処理の時間をモデル化できることを確認した。第二にブラックボックスアプローチで並列処理の時間をモデル化できることを確認した。第三にモデルパラメータに非負制約を付けて過剰適合抑制し、モデルの予測力が向上することを確認した。

第5章では終章として全体の総括を行った。

以上総じて、殆ど全ての並列処理においてその効率が評価できる性能評価指標を提案してその有効性を確認した。また並列処理時間のモデル化方法を、ホワイトボックスとブラックボックスという2つの対極的なアプローチから提案した。モデルの基底関数とモデル化に使った観測データが引き起こす過剰適合を抑制してモデルの予測精度を向上する、モデルパラメータ推定方法を提案して有効性を確認した。これらの並列処理の性能評価に関する成果は、高速計算を実現する並列計算機を有効利用する上で、現在と将来において極めて有意義な成果であると確信する。