

自動並列化コンパイラにおける条件付データフロー解析技術の評価

5W-05

高島 志泰^{1,2)}

1) アドバンスト並列化コンパイラ研究体

2) (株)日立製作所システム開発研究所

1. はじめに

並列マシンを用いてプログラムの実行を高速化するためには、プログラムからより多くの並列性を抽出することが重要である。このため、自動並列化コンパイラ WPP (Whole Program Parallelizer) [1]では、手続き間にまたがるデータ依存関係も解析してループを並列化する。しかし、ループ中の条件分岐により不明なデータ依存関係がある場合、原理的には並列化可能でも WPP では並列化できない。この問題を解決する方法として条件付データフロー解析技術[2]が提案されているが、コンパイラへの実装例はない。本稿では、この技術の WPP への実装と評価について述べる。

2. データフロー解析

2. 1 従来のデータフロー解析

ループに対する従来の配列データフロー解析[3]では、ループ内で参照される各配列に対して、以下の 4 種類の配列参照範囲を調べる。

- ・ USE：使用される可能性のある参照範囲
- ・ EUSE：定義される前に使用される可能性のある参照範囲
- ・ MOD：定義される可能性のある参照範囲
- ・ KILL：必ず定義される参照範囲

並列化可能性を判定するために配列参照範囲同士の重なりを調べる。つまり、図 1 のようにループ独立依存やプライベート化可否を判定する。図中の $MOD[I:I-1]$ は、I 番目のイタレーションにおける参照範囲、 $USE[I:I-1]$ は、1 から I-1 番目のイタレーションにおける参照範囲を表わす。

ループ独立依存判定：	
$(MOD[I:I-1] \cap EUSE[I]) = \phi$ 且つ	(1)
$(USE[I:I-1] \cap MOD[I]) = \phi$ 且つ	(2)
$(MOD[I:I-1] \cap MOD[I]) = \phi$	(3)
プライベート化可否判定：	
ループ独立依存判定が不成立 且つ	
$(MOD[I:I-1] \cap EUSE[I]) = \phi$	(4)

図 1：並列化可能性解析方法

2. 2 条件付データフロー解析

条件付データフロー解析は、配列の参照範囲を求めるときに、参照範囲とその範囲を参照するときの参照条件をペアにして解析する。図 2 に例を

示す。図 2 (a) のプログラム例中の配列 A に対する従来の解析では、図 1 (1) の等式が成立せず、並列化不可能となる(図 2 (b))。一方、参照条件付の参照範囲を用いると、図 1 (1) の等式は、 $X \leq 5$ のとき成立する(図 2 (c))。同様に図 1 (2) や (3) も計算することによって、並列化が可能となる条件式が算出できる。

この条件式の成否判定がコンパイル時にできる場合はコンパイル時に行ない、できない場合は実行時に行う必要がある。このため、ループをマルチバージョン化し、条件式の成否により並列ループ/逐次ループを選択するコードを生成する。

DO I=1,N IF (X>5) THEN A(I) = A(I-1) + C ENDIF B(I) = A(I) + D ENDDO	$MOD[I:I-1] = A[I:I-1]$ $EUSE[I] = A[I-1:I]$ $MOD[I:I-1] \cap EUSE[I]$ $= A[I:I-1] \cap A[I-1:I]$ $= A[I-1] \neq \phi$
---	--

(a) プログラム例

(b) 従来の解析手順

$MOD[I:I-1] = X > 5, A[I:I-1]$ $EUSE[I] = X > 5, A[I-1:I] \cup X \leq 5, A[I]$ $MOD[I:I-1] \cap EUSE[I]$ $= (X > 5, A[I:I-1]) \cap (X > 5, A[I-1:I] \cup X \leq 5, A[I])$ $= X > 5, A[I-1] \rightarrow X \leq 5, \phi$
--

(c) 参照条件付き参照範囲を用いた解析手順

図 2：適用例

3. 条件付データフロー解析の実装

条件付データフロー解析の実装では、WPP に以下の 3 処理を追加した(図 3、網掛け部分)。

- (1) 参照条件の参照範囲への付加
- (2) 並列化可能条件式の算出
- (3) マルチバージョン化

これらを以下で説明する。

- (1) 参照条件の参照範囲への付加：if 文の条件式を if ブロック内の各文の参照範囲に付加する。
- (2) 並列化可能条件式の算出：(1) で求めた参照範囲から算出する。ループ独立依存条件式は、図 1 (1) から (3) に対して図 2 (c) のように算出した 3 つの条件式の AND をとる。プライベート化可

能条件式は、同様に図1(4)に対する条件式から、ループ独立依存条件式を除いたものとする。

(3) マルチバージョン化：(2)で算出された条件式を用いて、プライベート化なし並列ループ、プライベート化あり並列ループ、及び、逐次ループの3つのうちの1つを実行時に選択するコードを生成する。プライベート化なし並列ループの実行可能条件は、各配列のループ独立依存条件式のANDとする。プライベート化あり並列ループの実行可能条件は、各配列のプライベート化可能条件式のANDとする。

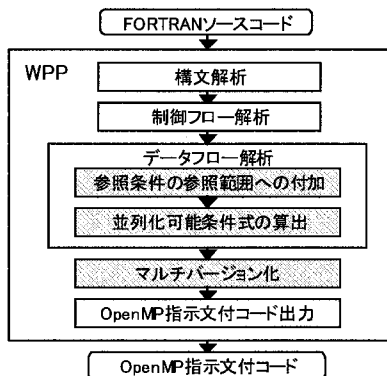


図3：WPPの構成

4. 評価

WPPに条件付データフロー解析を実装し、SPECfp95/apsi中のループ(図5(a)中の解析対象ループL)に適用した。ただし、WPPの配列参照範囲解析の制限により、呼び出し先での配列の寸法を定数に変更した。従来のWPPではSUB2の手続き内のIF文により配列Aのデータ依存関係が不明になり、並列化できなかった。

図5(b)は、条件付データフロー解析を適用した結果、出力されたOpenMPプログラムである。マルチバージョン化により、プライベート化あり(図中(1))となし(図中(2))の並列ループが生成され、実行時には(1)が選択された。

日立スーパーコンSR8000の1ノード(SMP)で実行した結果、8プロセッサで6.28倍の台数効果を得た。また、SGI社のDSMマシンOrigin2000(32プロセッサ)で実行した結果、Origin2000上では6プロセッサで2.06倍の台数効果を得た(図4)。

5. おわりに

WPPに条件付データフロー解析を実装し、SPECfp95/apsi中のループに適用してSR8000と

Origin2000上で評価を行い、2から6倍の台数効果を得た。

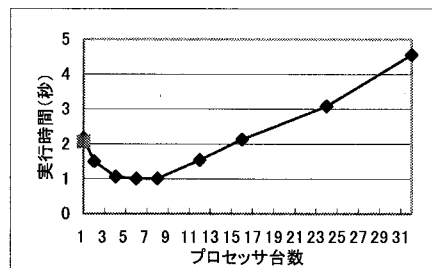
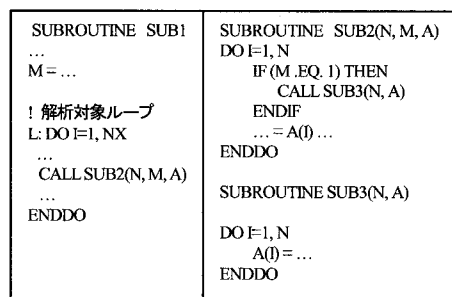
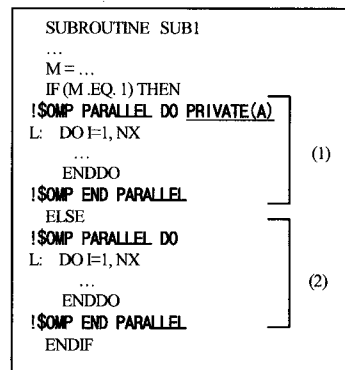


図4：Origin2000上での性能評価



(a) 解析対象ループの構造



(b) 出力結果

図5：apsiプログラム例

参考文献

- [1] 青木 雄一郎, 他, "手続き間コンパイラ WPP の試作 -実機性能評価-", 情報処理学会研究会報告, 98-ARC-72, pp.43-48, 1998.
- [2] Moon, S, et al., "Predicated Array Data-Flow Analysis for Run-time Parallelization", ICS'98, 1998, 204-211.
- [3] Hall, Mary W, et al., "Interprocedural Analysis for Parallelization", LCPC'95, 1995.