プログラミング言語 10－2
（1987. 2. 13）

多重ループ内の配列多重添字解析方式

石田　和久[*]　　　金田　泰[**]

[*] 日立製作所　ソフトウエア工場　[**] 同　中央研究所

ＦＯＲＴＲＡＮコンパイラにおけるベクトル化においては、配列の添字解析方式が、最も重要な位置を占める。配列の添字解析とは、プログラム中の２つの配列参照の添字を比較することにより、それらの同一アドレス参照の有無を明らかにすることである。従来の配列の添字解析方式においては、最内側ループを主な解析対象にしていた為、十分なベクトル化を行えないプログラムもあった。この報告では、多重ループを解析対象とする配列の添字解析方式について述べる。本方式により、ループ１重化や、ループ交換による多重ループのベクトル化を行えるようになる。

An analyzer of multi-dimensional array subscripts in nested loops

Kazuhisa Ishida[*]　**Yasusi Kanada**[**]

[*] Software Works,Hitachi Ltd,5030,Totuka,Totuka,Yokohama,Kanagawa,244 Japan

[**] Central Research Laboratory,Hitachi Ltd,1-2830,Higashikoigakubo,Kokubunji, Tokyo ,185 Japan

An analyzer of array subscripts plays an important part of a vectorizing for FORTRAN compiler. Where the purpose of an analyzer of array subscripts is to make it clear by comparison with two array subscripts whether two array references access the same address or not. The formal method analyse mainly the innermost loop, so there exists some programs that can not be vectorized.

We show a method to analyse an array sub-scripts in nested loops. By this method, we can vectorize nested loops by loop-exchange or loop-singlize.

１．はじめに

ＦＯＲＴＲＡＮコンパイラにおける最適化、特に、スーパーコンピュータ用コンパイラにおける自動ベクトル化においては、配列の添字解析方式が、最も重要な位置を占める。配列の添字解析とは、プログラム中の２つの配列参照の添字を比較することにより、それらの同一アドレス参照の可能性や、順序性を明らかにすることである。

最内側ループに対する配列の添字解析方式については、［１］、［２］などの研究があり、多重ループに対するものとしては、多重ループを１重化してベクトル化することを目的とした研究［３］がある。

この報告では、多重ループ中の２つの配列参照の添字が、複数のループのインダクション変数の線形結合で表されている場合（この添字を多重添字という）の配列の添字解析方式について述べる。

まず、第２章で、添字解析を行う準備として、規準インダクション変数と配列添字の標準形について述べる。第３章で、本方式の出力であるループ依存性情報について述べる。これは、２つの配列参照の同一アドレス参照の可能性の有無を、ループの繰返しに関して表したものである。

第４章と第５章で、配列添字の標準形を比較することにより、ループ依存性情報を求める方法について述べる。そして、第６章で結論を述べる。

２．規準インダクション変数と配列添字の標準形

配列の添字解析を行う準備として、インダクション変数と規準インダクション変数、及び、配列添字の標準化と標準形を以下のように定義する。

［定義］（インダクション変数）

ループ不変な値が、ループの繰返し毎に加算（ないし、減算）される変数をインダクション変数といい、このループ不変な値を増分値という。また、インダクション変数のループ１回目の値を、初期値という。

［定義］（規準インダクション変数）

ループγに対し、初期値０，増分値１，の仮想的なインダクション変数Ｉγを設け、これを、ループγの規準インダクション変数という。

［定義］（標準化と標準形）

配列添字Ｊが、上記規準インダクション変数により、

$$J = C_0 + \sum_{\gamma:\text{ループ}} C_\gamma I_\gamma \qquad (1)$$

$$(C_0,\ C_\gamma:\text{定数})$$

と表されるとき、添字Ｊは、標準化可能であるといい、（１）の右辺をＪの標準形とよぶ。

また、標準形を求めることを標準化と呼ぶ。

配列添字を標準化することにより、配列添字のループ繰返しにおける変化を、コンパイル時に予測し易くなる。

図１を用い配列添字の標準形の例をしめす。
配列参照②，⑤，⑦における添字の標準形表現はそれぞれ

$4I\gamma_1$，$2+8I\gamma_1$，$7+16I\gamma_1+3I\gamma_2$

となる。

標準化処理では、次のような情報ももとめており、従来より強力な添字解析を実施している。

（ａ）ループγ_1におけるインダクション変数ｊの定義は２箇所ある。：③と⑧

（ｂ）インダクション変数ｉの初期値はループγ_1のインダクション変数ｊである。：④

```
         j = 0                        ①
         DO 10 k1 = 1, N              (ループγ1)
           A(j) =....                 ②
           j = j+1                    ③
           i = j                      ④
           DO 20 k2 = 1, N            (ループγ2)
             m = 2*j
             A(m) =....               ⑤
             i = i+1                  ⑥
             A(3*i+j)=.....           ⑦
20         CONTINUE
           j = j+3                    ⑧
10       CONTINUE
```

図1　規準インダクション変数による添字標準化

３．ループ依存性情報

ベクトル化などの最適化を行なう場合、ループ内の２つの配列参照が、

（ａ）ループの同一の繰り返しにおいて、同一アドレスを参照するか否か。

（ｂ）ループの異なる繰り返しにおいて、同一アドレスを参照するか否か。

という情報が重要である。　これらの情報をループ依存性情報と呼ぶ。ループ依存性情報とは、２つの配列参照の同一アドレスの参照が、「必ず生じる」のか、あるいは、「生じる可能性がある」のか、及び、それが「ループの何回の繰り返し後において生じる」のか、についての性質であり、正確には、以下の（１），(２），(３）の３つの性質からなる。　ここに、γはループ、Ｕ、Ｖは、γ内の配列の参照を表わし、ｎは ループγの何回の繰り返し後に、ＵとＶが同一アドレスを参照するのかを表わす負でない整数である。

（１）ＵとＶのγに関するループｎ回完全一致性。

これは、「Ｕが参照したアドレスを、Ｖが ループγのｎ回後の繰り返しにおいて 必ず参照する」という性質である。

（２）ＵとＶのγに関するループｎ回一致可能性。

これは、「Ｕが参照したアドレスを、Ｖが ループγのｎ回の繰り返し後に 参照する可能性がある」 という性質である。逆に言えば、「ｎ回後の繰り返しにおいて、必ず異なるアドレスを参照するとは言えない」という性質である。

（３）ＵとＶのγに関するループｎ回以上一致可能性。

これは、「Ｕが参照したアドレスを、Ｖが ループγのｎ回以上の繰り返し後に 参照する可能性がある」 という性質である。逆に言えば、「ｎ回以上のすべての繰り返しにおいて、必ず異なるアドレスを参照するとは言えない」という性質である。

図２に、上記情報の分類図を示す。

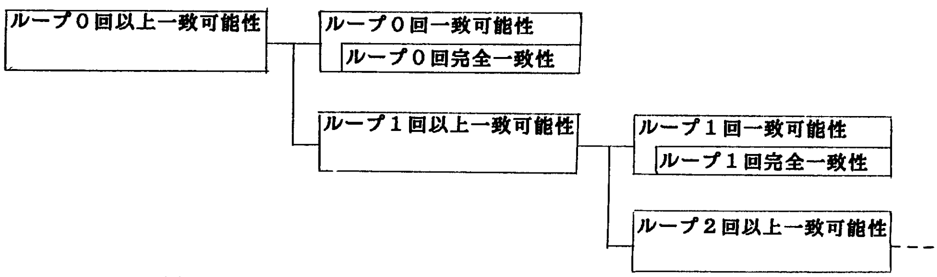

図２．ループ依存性情報の分類図

図３に、上記情報の例を示す。図３－Ａ　に対しては、Ｕ1で参照したアドレスを、Ｖ1が ループγ_1の１回の繰り返し（次の繰り返し）で 必ず参照するから、Ｕ1とＶ1の γ_1に関するループ１回完全一致性が有る。　図３－Ｂ　では、Ｕ2で参照したアドレスを、Ｖ2が ループγ_2の０回の繰り返し（同じ繰り返し）において、ｊ＝１ なら参照し、ｊ＞１ なら参照しない。従って、Ｕ2とＶ2の γ_2に関するループ０回完全一致性は無く、ループ０回一致可能性が有る。　図３－Ｃ　では、ｊ$\geqq$１ であるから、Ｕ3で参照したアドレスをＶ3がループγ_3の０回以上の繰り返し後において参照することはない。従って、Ｕ3とＶ3の γ_3に関するループ０回以上一致可能性は無い。

```
      DO 10 i = 1, N            (ループγ1)
        A(i) =...               (U1)

10      ...= A(i-1)             (V1)
```

Ａ　ループ１回完全一致性が有る例

```
      DO 20 i = 1, N            (ループγ2)
        A(i) =...               (U2)
        DO 20 j = 1, N
20        ...= A(i+j-1)         (V2)
```

Ｂ　ループ０回一致可能性が有る例

```
      DO 30 i = 1, N            (ループγ3)
        A(i) =...               (U3)
        DO 30 j = 1, N
30        ...= A(i+j)           (V3)
```

Ｃ　ループ０回以上一致可能性が無い例

図３．ループ依存性情報

４.標準形の係数の比較による添字比較方法

ループ依存性情報は、配列添字の標準形の係数を比較して求める。

ループγ'内の配列の2つの参照ＵとＶの添字がそれぞれ、

$$C_0 + \sum_{\gamma:\text{ループ}} C\gamma I\gamma$$

$$D_0 + \sum_{\gamma:\text{ループ}} D\gamma I\gamma$$

（C_0，$C\gamma$，D_0，$D\gamma$：定数）

と標準化されているとき、０以上の整数ｎに対して、

（１）次の(1-1)、(1-2)、(1-3) 全てが成立すれば、ＵとＶの γ' に関するループｎ回完全一致性が有る。

(1-1) $C_0 = D_0 + D\gamma' * n$

(1-2) $C\gamma = D\gamma$　　　（γ：γ'と等しいループ 又はγ'の外側ループ）

(1-3) $C\gamma = D\gamma = 0$　　（γ：上記以外のループ）

（２）次の (2-1)、または、(2-2)が 成立すれば、ＵとＶの γ' に関するループｎ回一致可能性は無い。

(2-1) $C_0 > D_0 + D\gamma' * n$ かつ $C\gamma \geq D\gamma$ かつ　$C\gamma'' * D\gamma'' = 0$

（γ：全てのループ；　γ''：γ'と異なり、かつ γ'の外側ループでないループ）

(2-2) $C_0 < D_0 + D\gamma' * n$ かつ $C\gamma \leq D\gamma$ かつ　$C\gamma'' * D\gamma'' = 0$

（γ：全てのループ；　γ''：γ'と異なり、かつ γ'の外側ループでないループ）

（３）次の　(3-1)、または、(3-2)が 成立すれば、ＵとＶの γ' に関するループｎ回以上一致可能性は無い。

(3-1) $C_0 > D_0 + D\gamma' * n$ かつ $D\gamma' \leq 0$ かつ　$C\gamma \geq D\gamma$ かつ $C\gamma'' * D\gamma'' = 0$

（γ：全てのループ；　γ''：γ'と異なり、かつ γ'の外側ループでないループ）

(3-2) $C_0 < D_0 + D\gamma' * n$ かつ $D\gamma' \geq 0$ かつ　$C\gamma \leq D\gamma$ かつ $C\gamma'' * D\gamma'' = 0$

（γ：全てのループ；　γ''：γ'と異なり、かつ γ'の外側ループでないループ）

ベクトル化などに有用な性質は、完全一致性があること、および、一致可能性が 無いことである。多くの場合、上記 (1),(2),(3)によりループ依存性情報を求めることができるが、求められない場合には、２つの添字の標準形が等しいとして 整数不定方程式をたて、その解の挙動を調べることによって ループ依存性情報を求める。これについては、次章で述べる。

図４のプログラムを例として、添字比較方法を説明する。
配列Ａの参照Ｕ，Ｖの添字の標準形はそれぞれ、

$(1+I\gamma_1,\ 2+I\gamma_1+I\gamma_2)$，

$(2+I\gamma_1+I\gamma_3,\ 1+I\gamma_1)$

となる。１次元目どうし、及び、２次元目どうしに、上記（３）の（3-2）を適用すれば、ＵとＶの γ_1 及び γ_2 に関するループ０回以上一致可能性が、ともに無く、ＶとＵの γ_1 及び γ_2 に関するループ１回以上一致可能性が、ともに無いことがわかる。

```
      DO 10 k = 1, N             (ループγ1)
        DO 20 i = k+1, N         (ループγ2)
          A(k,i) = T             (U)
          DO 30 j = k+1, N       (ループγ3)
30          S = A(j,k)           (V)
20      CONTINUE
10    CONTINUE
```

図４　添字比較の例

５．整数不定方程式による添字比較方法

本章では、第４章で述べた方法によりループ依存性情報が求められなかった場合、さらに詳しく配列添字を解析することによって、ループ依存性情報を求める方法について述べる。

ループループ γ_0 内の、Ｎ次元配列の２つの参照Ｕ，Ｖの添字を　それぞれ、

$(I1,\ I2,\ \cdots,\ IN)$

$(J1,\ J2,\ \cdots,\ JN)$

とし、各Ｉi，Ｊiは、

$$Ii = Ci0 + \sum_{\gamma:\text{ループ}} Ci\gamma\, I\gamma$$

$$Ji = Di0 + \sum_{\gamma:\text{ループ}} Di\gamma\, I\gamma$$

$(Ci0,\ Di0,\ Ci\gamma,\ Di\gamma$:定数$)$

と標準化されているものと仮定する。

ループの集合　Ω，Θを、

$\Omega = \{\gamma:$ループ $\mid \gamma$はγ_0 の内側ループ$\}$

$\Theta = \Omega \cup \{\gamma_0\}$

とおき、式ＪiΩ，ＪiΘ　を

$$Ji\Omega = Di0 + \sum_{\gamma\in\Omega} Di\gamma\, I\gamma' + \sum_{\gamma\notin\Omega} Di\gamma\, I\gamma$$

$$Ji\Theta = Di0 + \sum_{\gamma\in\Theta} Di\gamma\, I\gamma' + \sum_{\gamma\notin\Theta} Di\gamma\, I\gamma$$

とおく。

ここで、γ_0に関するループ n回一致可能性を調べるには、ループγ_0の内側ループγの規準インダクション変数は、UとVとでは独立に変化するものとして扱う必要があるため、Vに対するものを、$I\gamma'$と改名してある。（JiΩ）　更に、ループn回以上一致可能性を調べるには、ループγ_0の規準インダクション変数も　UとVとでは、独立に変化するものとして扱う必要があるため、Vに対するものを、$I\gamma'_0$　と、改名してある。（JiΘ）

このとき、負でない整数nに対して、次の十分条件（3）、（4）が成立する。

（3）（ループn回一致可能性が無いための十分条件）

次の連立整数不定方程式が解をもたなければ、UとVのγ_0に関するループn回一致可能性は無い。

$Ii = Di0 * n + Ji\Omega$　　(i = 1, ・・・, N)

（4）（ループn回以上一致可能性が無いための十分条件）

連立整数不定方程式

$Ii = Di0 * n + Ji\Theta$　　(i = 1, ・・・, N)

に対して、次の(4-1)、または(4-2)が成立すれば、UとVのγ_0 に関するループn回以上一致可能性は無い。

(4-1)　解をもたない

(4-2)　解空間上で常に$I\gamma_0 > I\gamma'_0$

上記（3）、（4）の十分条件を調べるために、まず規準インダクション変数の範囲を考慮せずに、連立方程式の解を補助変数の線形結合として代数的に求めておく。その結果と、各規準インダクション変数が、0以上　{（ループの繰返し回数）－1} 以下であることを利用して解空間を整数範囲で近似し、近似した解空間上で、十分条件（3）、（4）を調べる。

図5のプログラムを例として説明する。

```
      DO 10 i=1,10                      (ループγ1)
        DO 20 j=1,10                    (ループγ2)
 20       A(i+3j-5) =  ・・・           (U)
        DO 30 k=1,10                    (ループγ3)
 30       ・・・      =A(-i-4k+5)       (V)
 10 CONTINUE
```

図5．整数不定方程式による添字比較の例

参照U，Vの添字の標準形はそれぞれ

$-1 + I\gamma_1 + 3I\gamma_2$、$-2I\gamma_1 - 4I\gamma_3$

となる。

UとVのγ_1 に関するループ0回一致可能性 の有無は、4章の十分条件（2）を用いて決定することはできないため、上記（3）の十分条件を調べる。解析すべき方程式は、

$-1 + I\gamma_1 + 3I\gamma_2 \;=\; -2I\gamma_1 - 4I\gamma'_3$ ・・・・①

である。この方程式の解は、補助変数t_1，t_2を用いて、

$I\gamma_1 = 3t_1 - 2t_2 - 1$　　・・・・・②

$I\gamma_2 = -2t_1 + 1$　　・・・・・③

$I\gamma'_3 = t_2$　　・・・・・④

を満たす。

ループγ_1，γ_2，γ_3のループ繰返し回数が１０であるから、
規準インダクション変数 $I\gamma_1, I\gamma_2, I\gamma_3'$ は不等式

$0 \leq I\gamma_1 \leq 9$ ・・・・・⑤

$0 \leq I\gamma_2 \leq 9$ ・・・・・⑥

$0 \leq I\gamma_3' \leq 9$ ・・・・・⑦

を満たす。

③と⑥および補助変数t_1が整数であることから、

$1 \leq I\gamma_2 \leq 9$ ・・・・・⑧

を得る。よって、⑤、⑦、⑧ より、

（①の左辺）－（②の右辺）$= 2I\gamma_1 + 3I\gamma_2 + 4I\gamma_3' - 1 \geq 2 > 0$

従って、方程式　は近似した解空間⑤、⑦、⑧上で解をもたない。よって、ＵとＶのγ_1に関するループ０回一致可能性は無い。

６．終りに

以上、多重ループ内の配列多重添字解析方式として、配列参照の添字の標準形と、それを比較してループ依存性情報を求める方法について説明した。本方式により、従来より一般的かつ強力な配列添字解析が可能となる。その結果、ＦＯＲＴＲＡＮコンパイラにおけるベクトル化率の向上や、その他の配列に対する最適化の強化を図ることが可能となる。

参考文献

［１］　高貫　他：Ｓｏｍｅ　Ｃｏｍｐｉｌｉｎｇ　Ａｌｇｏｒｉｔｈｍｓ　ｆｏｒ ａｎ　Ａｒｒａｙ　Ｐｒｏｃｅｓｓｏｒ，Ｐｒｏｃ．　ｏｆ　３ｒｄ　ＵＳＡ－Ｊａｐａｎ　Ｃｏｍｐｕｔｅｒ　Ｃｏｎｆ，ｐ２７３－２７９　（１９７８）．

［２］　梅谷　他：内蔵ベクトル演算機能のための自動ベクトルコンパイラ方式，情報処理，Ｖｏｌ．２４，Ｎｏ．２　ｐ２３８－ｐ２４８　（１９８３）．

［３］　津田　他：ループ間にまたがるデータ参照関係をもつ多重ループのベクトル化，情報処理，Ｖｏｌ．２６，Ｎｏ．３　ｐ５３６－ｐ５４４　（１９８５）．

［４］　金田　他：Ｆｏｒｔｒａｎ最適化の強化－大域配列データフロー解析法－，情報処理学会第３２回全国大会，４Ｆ－３