

# 行政院國家科學委員會補助專題研究計畫成果報告

## 平行編譯環境中利用切割技術來最大化不規則相依 迴圈平行度方法之探討

計畫類別：□個別型計畫      ½整合型計畫

計畫編號：NSC89 - 2213 - E - 009 - 200

執行期間： 89年8月1日至90年7月31日

計畫主持人：陳正

共同主持人：

本成果報告包括以下應繳交之附件：

赴國外出差或研習心得報告一份

赴大陸地區出差或研習心得報告一份

出席國際學術會議心得報告及發表之論文各一份

國際合作研究計畫國外研究報告書一份

執行單位：國立交通大學資訊工程學系

中 華 民 國    90 年   10 月 29 日

# 行政院國家科學委員會專題研究計畫成果報告

## 平行編譯環境中利用切割技術來最大化不規則相依迴圈平行度方法之探討

### A Study of Partitioning Techniques to Maximize Parallelism of Nested Loop with Non-uniform Dependences in Parallelizing Compiler Environment

計畫編號：NSC 89-2213-E-009-200

執行期限：89年8月1日至90年7月31日

主持人：陳正 國立交通大學資訊工程學系

計畫參與人員：李宜軒、蔡明龍、張雁翔

國立交通大學資訊工程學系

#### 一、中英文摘要

本年度的計畫主要是擴展舊有計畫所開發出的理論，提出新的理論，在平行編譯系統中利用切割技術來平行化不規則相依迴圈。我們也將新提出的方法建立在 SUIF 編譯套件及以往發展的 SEESMA 模擬評估環境之中，評估其執行效能。根據我們的評估結果，新提出的方法可以有效開發不規則相依迴圈的平行度，降低整體執行時間。最後並提出演算法將我們提出的所有不規則相依迴圈平行化方法做一個整合，針對不同型態的不規則相依迴圈，選擇最適當的平行化方法。

**關鍵詞：** 不規則相依迴圈、迴圈切割、迴圈排程、迴圈平行化

#### Abstract

This project extends our earlier method and proposes a new method to maximize parallelism of nested loops with non-uniform dependences in parallelizing compiler environment. We also implement our methods on SUIF compiler and evaluate their execution performance on SEESMA. Based on our evaluation results, our methods can explore parallelism of nested loops effectively and decrease the entire execution time. Finally, we propose an algorithm to

integrate all methods we have been proposed, which can select the most appropriate parallelization method according to every styles of non-uniform nested loops.

**Keywords:** *Non-uniform Dependence Loop, Loop Partitioning, Loop Scheduling, Loop Parallelization*

#### 二、計畫緣由與目的

近幾年來，資料相依 (data dependence) 的問題已被廣泛的討論。資料相依的狀況主要可分為兩種型式，一為規則性的資料相依 (uniform/regular data dependence)，另一為不規則資料相依 (non-uniform/irregular data dependence)。規則性資料相依由於迴圈中的每一個迴圈元素中都是固定的，而且其迴圈形式 (loop pattern) 較簡單，相關方面的研究已經非常純熟[1-2]。不但有完整且有效的理論基礎，大部分也已實現在各種不同的平行編譯環境中；然而在實際的標竿程式中，仍存在不少不規則相依迴圈[3]。

以往在國外關於不規則相依迴圈的平行化方法的研究，較具代表性的主要分成三大類：相依則化技巧 (Dependence Uniformization Technique) 及其延伸方法 [4-5]、最小相依距離方塊化 (Minimum Dependence Distance Tiling)[6-7] 以及三區

域分割法 (Three Region Partitioning Technique)[8-9]; 國內相關的研究則有最小相依向量集 (Minimal Dependence Vector Set)[10] 及有效規則化迴圈執行方法的提出[11]。這些方法通常不能有效開發不規則相依迴圈的平行度, 而且執行時間複雜度往往偏高。另外, 我們也在國科會支持的整合型計畫下, 致力於有關方面的研究。我們提出的方法主要可分成四大類: 平行部分切割 (Parallelization Part Splitting, PPS)、部分平行化之個別處理 (Partial Parallelization Decomposition, PPD)、不規則迴圈維度交換 (Irregular Loop Interchange, ILI) 以及增長型相依型態偵測 (Growing Pattern Detection, GPD)[12]。針對三域切割方面, 我們也提出最佳化三區域切割法 (OTRP Method) [13], 改善原本三區域切割法的效能, 開發更多的平行度。

除了理論的研究, 我們也將這些方法實作於 Stanford 大學發展的 SUIF 平行編譯套件中[14], 並在 SEESMA 模擬評估環境之下評估執行效能[15]。根據理論研究及模擬評估的結果, 雖然 OTRP 及我們提出的其他方法可以達到不錯的結果, 但是仍然有未開發的平行度存在於不規則相依迴圈之中。因此本年度的計畫目標, 是要延伸最佳化三區域切割, 並提出新的理論及方法, 開發這些不規則相依迴圈中潛在的平行度; 同時也會將新的方法實作於 SUIF, 讓整個平行編譯環境更趨完整。最後再將我們到目前為止發展出來的所有最大化不規則相依迴圈平行度的方法加以整合, 提出一個演算法, 針對各種型態的不規則相依迴圈, 選擇最適合的平行化方法。有關新提出的方法以及模擬評估的結果, 將分別敘述如下。

### 三、結果與討論

在本計畫中共提出二個在平行編譯環境中最大化不規則相依迴圈平行度的方法, 分別實做於 SUIF 平行編譯套件並用 SEESMA 加以評估。一個名為兩階段切割方式 (Two-Stage Partitioning, TSP), 它實際上是延伸之前的最佳化三區域切割

(OTRP) 所發展的兩階段式不規則相依迴圈切割方式, 針對 OTRP 所留下的循序執行區塊做平行化切割以開發更多的平行度。TSP 第一階段與 OTRP 相同, 會將迴圈空間分成  $Area_1$ ,  $Area_2$  及  $Area_3$  等三個區域:  $Area_1$  屬於反向相依開頭 (anti-dependence head) 及正向相依結尾 (flow dependence tail) 的聯集, 但是不包含在正向相依開頭 (flow dependence head) 中的迴圈元素;  $Area_2$  屬於反向相依開頭及正向相依開頭的聯集;  $Area_3$  包含剩下的迴圈元素。根據 OTRP 的分析, 經過複製重新命名 (copy-renaming) 的方法消除反向相依的關係之後, 所有在  $Area_1$  和  $Area_2$  的迴圈元素都可以被平行執行,  $Area_3$  則需要循序執行, 整體執行順序為  $Area_1$   $Area_3$   $Area_2$ 。OTRP 已證明其切割出來的  $Area_3$  面積比其他三區域切割方法要小, 因此可以達到最佳的效果; 至於  $Area_3$  必須循序執行的問題, 則在 TSP 第二階段中加以解決。第二階段的主要步驟是在  $Area_3$  中找到 lexicographical 執行順序最前面的正向相依開頭迴圈元素  $(x, y)$ , 而且它對應的結尾亦屬於  $Area_3$ 。因為不存在相依關係, 所有在  $Area_3$  中而其 lexicographical 執行順序小於  $(x, y)$  的迴圈元素可以切成一個平行區, 其餘在  $Area_3$  中的迴圈元素亦可重複同樣的步驟。然而如果相關頭對應的結尾位於之前的分割區, 則此相依關係不存在。TSP 第二階段方法有以下五個特性:

1. 巢狀迴圈中如果所有相依開頭對應的結尾都屬於之前的分割區或  $Area_1$ , 則在目前分割區的起始節點及最後相依開頭所屬節點之間, 符合 lexicographical 執行順序的所有迴圈元素都可以被平行執行。
2. 在同一個由第二階段 TSP 方法所得到的分割區中的迴圈元素可以被平行執行。
3. 巢狀迴圈中因為要符合 lexicographical 執行順序, 所以在一個由第二階段 TSP 方法所得到的分割區中的迴圈元素數目是最大值。
4. 分割過程中在檢查相依關係時, 只需檢查在  $Area_3$  之中的迴圈元素。
5. 迴圈元素  $i, j$  分別為正向相依的開頭及結尾, 若要檢查  $i, j$  之間的相依關係,

只要檢查 lexicographical 執行順序在  $j$  之後的迴圈元素即可。

根據以上特性，我們可以使用改良式整數程式技術 (integer programming technique) 來計算相依開頭迴圈元素，降低演算法複雜度，提升整體執行效能。

TSP 演算法我們已經成功建立在 SUIF 平行編譯套件並用 SEESMA 加以評估，使用的標竿程式為四個規格模組 (standard model)[4, 16] 及選自 [16-17] 的二個實際程式片段。根據評估結果，TSP 在所有標竿程式中，都可以達到比其他方法更佳的執行時間及增益 (speedup)，由此可見 TSP 確實是一個有效開發不規則相依迴圈平行度的方法。圖 1 為 TSP 與其它方法在 SEESMA 上模擬的增益評估圖，使用標竿程式為[18]中的實例，其餘的評估結果可參考[18-19]。TSP 的研究成果已發表於會議論文[19]。

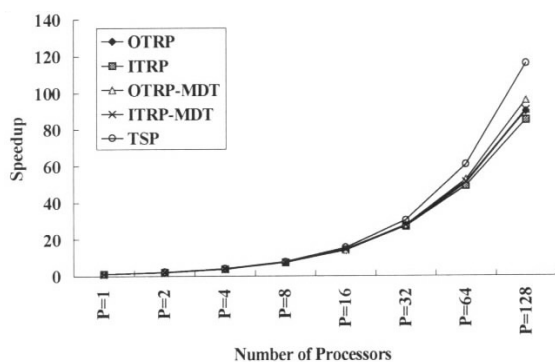


圖 1 [18]實例在 SEESMA 中的增益評估

本計畫提出的另一個方法名為最佳化相依非矩形區塊切割法 (Optimized Dependence Convex Hull Partitioning, ODCHP)，它以非矩形理論 (Convex Hull theory) 為設計基礎，將迴圈空間分成多個不同大小的平行區，並可與其他不規則相依迴圈平行化技術相結合，來提高其平行化程度。ODCHP 的設計理念和執行方法與 TSP 第二階段類似，差別在於 TSP 第二階段僅針對 OTRP 留下的循序執行區  $Area_3$  做平行化切割，而 ODCHP 則切割整個迴圈空間。ODCHP 的執行方法為在迴圈空間範圍中，找一個 lexicographical 執行順序在最前面的相依開頭迴圈元素  $(x, y)$ ，由於不

存在相依關係，所有 lexicographical 執行順序在  $(x, y)$  之前的迴圈元素可以被切割成一個平行區，其餘的迴圈元素亦可重複同樣的切割步驟。然而如果相依開頭對應的結尾位於之前的分割區，則此相依關係不存在。詳細定義如下：在巢狀迴圈中，由 ODCHP 切割方法所得到的第  $k$  個切割區，是從前一個切割區的邊界開始，直到一個相依開的前一個節點，而且此相依開頭所對應的相依結尾在此切割區中。ODCHP 同樣有以下五個特性：

1. (包含特性) 在巢狀迴圈中，如果所有相依開頭對應的相依結尾都屬於之前的 ODCHP 切割區，則所有在目前 ODCHP 切割區的起始節點及最後相依開頭所屬節點之間，符合 lexicographical 執行順序的迴圈元素，都可以被平行執行。
2. 在巢狀迴圈中，在同一個 ODCHP 切割區中的迴圈元素可以被平行執行。
3. 巢狀迴圈中因為要符合 lexicographical 執行順序，所以在一個 ODCHP 切割區中的迴圈元素數目是最大值。
4. (邊界特性) 根據 OTRP 的分析，正向相依開頭或結尾迴圈元素只會出現在特定的集合之中。因此，如果我們想要搜尋正向相依開頭或結尾，只要搜尋特定集合即可，不需搜尋整個迴圈空間，如此可以明顯減少時間複雜度。
5. (傳遞特性) 迴圈元素  $i, j$  分別為正向相依的開頭及結尾，若要檢查  $i, j$  之間的相依關係，只要檢查 lexicographical 執行順序在  $j$  之後的迴圈元素即可。

根據以上特性，與 TSP 相同，我們也可以使用改良式整數程式技術 (integer programming technique) 來計算相依開頭迴圈元素，降低演算法複雜度，提升整體執行效能。

ODCHP 演算法我們已經成功建立在 SUIF 平行編譯套件並用 SEESMA 加以評估，使用的標竿程式與 TSP 相同。根據評估結果，ODCHP 在所有標竿程式中，都可以達到比其他方法更佳的執行時間及增益 (speedup)，由此可見 ODCHP 同樣是一個有效開發不規則相依迴圈平行度的方法。

圖 2 為 ODCHP 與其它方法在 SEESMA 上模擬的增益評估圖，使用標竿程式為[4, 16]中的第一個規格模組，其餘的評估結果可參考[18, 20-21]。ODCHP 的研究成果已發表於會議論文[20]並被期刊論文接受[21]。

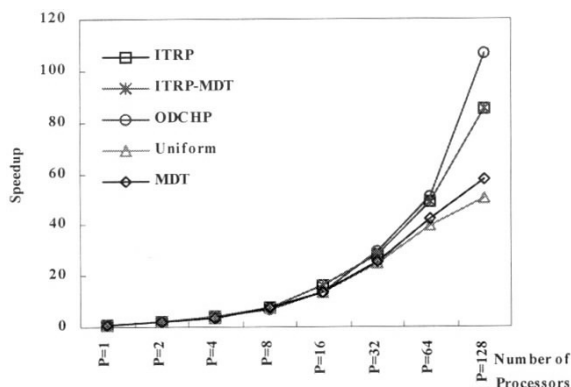


圖 2 Model 1 在 SEESMA 中的增益評估

最後再將我們到目前為止發展出來的所有最大化不規則相依迴圈平行度的方法加以整合，提出一個演算法，針對各種型態的不規則相依迴圈，選擇最適合的平行化方法。整合演算法及結果可參考[18]，並已發表於會議論文[22]。

#### 四、計畫結果自評

在本年度的計畫中，我們針對不規則相依迴圈，延伸過去的計畫成果，進一步探討其平行化的議題，並實做於 SUIF 平行編譯套件中，用 SEESMA 模擬評估環境評估結果。總體而言，本計畫大致達成以下幾點目標：

1. 延伸最佳三區域切割法 (OTRP)，發展出兩階段切割方式 (TSP)，解決 OTRP 留下的循序執行部分，儘量開發不規則相依迴圈潛在的平行度。
2. 運用 OTRP 的概念加上整數程式技術，發展出最佳化相依方塊切割方法 (ODCHP)，將整個迴圈空間分成數個可以平行執行的區塊。
3. 將 TSP 及 ODCHP 實作於 SUIF 平行編譯套件，並在自行發展的 SEESMA 模擬評估環境中評估執行效能。
4. 與現有平行化方法做比較，證明 TSP 及 ODCHP 均可達到高的執行效能。
5. 提出整合演算法，針對各種型態的不規

則相依迴圈，選擇最適合的平行化方式。

由以上幾點可知，本計畫確實能將最大化不規則相依迴圈平行度的議題做詳細的探討，延伸之前的方法改善其不足之處，並嘗試提出新的方法達到理想的效果。除了理論的研究，我們也將之實做並模擬評估其執行效能。整體看來，我們提出的方法大致都能達到預期的效果，有效開發不規則相依迴圈的平行度，並將目前所有提出的方法加以整合。所有的研究成果，均已發表於相關期刊及會議論文。

#### 五、參考文獻

- [1] Michael E. Wolf and Monica S. Lam, "A Loop Transformation Theory and an Algorithm to Maximize Parallelism", *Proc. of 3<sup>rd</sup> Workshop on Programming Languages and Compilers for Parallel Computing*, April 1990.
- [2] Michael Wolfe, **High Performance Compilers for Parallel Computing**, Addison-Wesley Publishing Company, 1996.
- [3] Zhiyu Shen et al. "An Empirical Study on Array Subscripts and Data Dependencies", *Proc. of International Conference of Parallel Processing*, pp. 145-152, 1989.
- [4] Ten H. Tzen and Loinel M. Ni, "Dependence Uniformization: A Loop Parallelization Technique", *IEEE Transactions on Parallel and Distributed System*, Vol. 4, pp. 547-558, May 1993.
- [5] Weijia Shang, Edin Hodzic, and Zhigang Chen, "On Uniformization of Affine Dependence Algorithms", *IEEE Trans. on Computers*, Vol. 45, No. 7, pp. 827-840, July 1996.
- [6] Shaw-Yen Tseng, Chung-Ta King, and Chuan-Yi Tang, "Minimum Dependence Vector Set: A New Compiler Technique for Enhancing Loop Parallelism", *Proc. of International Conference of Parallel and Distributed Systems*, pp. 340-346, Dec. 1992.
- [7] Swamy Punyamurtula and Vipin Chaudhary, "Minimum Dependence

- Distance Tiling of Nested Loops with Non-uniform Dependences”, *Proc. of 6<sup>th</sup> IEEE Symposium on Parallel and Distributed*, pp. 74-81, May 1994.
- [8] Zaafrani, A. and Ito. M., “Parallel Region Execution of Loops with Irregular Dependences”, *Proc. of the International Conference on Parallel Processing*, pp. 11-19, 1994.
- [9] Cho, C. K. and Lee, M. H., “A Loop Parallelization Method for Nested Loops with Non-uniform Dependences”, *Proc. of International Conference on Parallel Processing*, pp. 314-321, 1997.
- [10] Shaw-Yen Tseng, Chung-Ta King, and Chuan-Yi Tang, “Profiling Dependence Vectors for Loop Parallelization”, *Proc. of IPPS*, pp. 23-27, 1996.
- [11] 陳正, 吳全臨, “可延伸性多處理機系統架構設計之研究, 子計畫二: 平行編譯技術及其環境之探討”, 國科會期末報告, 1996年12月.
- [12] 平德林, 陳正, “一個有效率的不規則相依迴圈平行化技術”, 全國計算機會議論文集, 第三冊, 1997年12月.
- [13] Der-Lin Pean and Cheng Chen, “An Optimized Loop Partition Technique for Maximize Parallelism of Nested Loops with Non-uniform Dependences”, *Journal of Information Science and Engineering*, 1999.
- [14] Robert Wilson, Monica Lam, and John Hennessy et al., “An Overview of the SUIF Compiler System”, Computer Systems Lab Stanford University, 1996.
- [15] Chao-Chin Wu, Der-LinPean, Jen-Pin Sy, Jia-Rong Wu, Hsuan-Woei Huang, Jun-Long Huang, Jia-Lin Lee, Huey-Ting Hua, and Cheng Chen, “SEESMA: A Simulation and Evaluation Environment for Shared Memory Multiprocessor Architecture”, “*Proc. of NSC- Par A: Physical Science and Engineering*, Vol. 22, No. 4, pp. 524-538, July 1998.
- [16] Swarztrauber, P. A. and Sweet, R. A., “Effective Fortran Subprograms for the Solution of separable Elliptic Partial Differential Equations”, *ACM Trans. on Mathematical Software*, Vol. 5, No. 3, pp. 352-364, 1979.
- [17] Dongarra, J. J., Moler, C. B., Bunch, J. R., and Stewart G. W., *LINPACK Users' Guide by SIAM*, Philadelphia, 1979.
- [18] 平德林, “最大化不規則相依迴圈平行度的有效切割方式”, 國立交通大學研究所博士論文.
- [19] Der-Lin Pean and Cheng Chen, “A Combined Partitioning Technique to Maximize Parallelism of Nested Loops with Non-uniform Dependences”, *Proc. of 6<sup>th</sup> Workshop on Compiler Techniques for High-Performance Computing*, pp. 45-54, March 2000.
- [20] Der-Lin Pean, Guan-Joe Lai, and Cheng Chen, “An Optimized Dependence Convex Hull Partitioning Technique to Maximize Parallelism of Nested Loops with Non-uniform Dependence”, *Proc. of International Conference on Parallel and Distributed Systems*, pp. 367-374, July 2000.
- [21] Der-Lin Pean and Cheng Chen, “ODCHP: A New Effective Mechanism to Maximize Parallelism of Nested Loops with Non-uniform Dependences”, accepted and to appear in the *Journal of Systems and Software*.
- [22] Der-Lin Pean, Guan-Joe Lai, and Cheng Chen, “A Systematical Partitioning Mechanism for Nested Loops with Non-uniform Dependences”, *Proc. of International Computer Symposium*, pp. 82-90, Dec. 2000.