

## GFT中的向量编译技术

杨 桃 栏

(计算机科学系)

**摘 要** 本文介绍了一个高效而简洁的编译程序GFT及其所使用的主要编译技术。

**关键词** 编译, 向量, 表格

### 1. 引 言

向量机是为适应现代科学技术某些领域(如地质勘探、中长期天气预报及原子能研究等)的高速计算要求而产生的。在这些领域里,高速计算是解决问题的关键。一句(10天)的天气预报,如果用普通计算机运算,要10天以上才有结果。那么“预报”就会变成“汇报”,显然毫无意义。

获得高速的因素是多环节的。基本的是提供高速硬件(中央处理机及与其匹配的外围系统)。但高速的真正获得,则是软件各环节的任务,如研制适合向量机的算法。软件中每一个用户都要与之打交道的是系统软件,如操作系统、编译程序等。

程序语言的设计是一个重要环节<sup>[1]</sup>,而程序语言的编译则是其后的又一个重要环节。

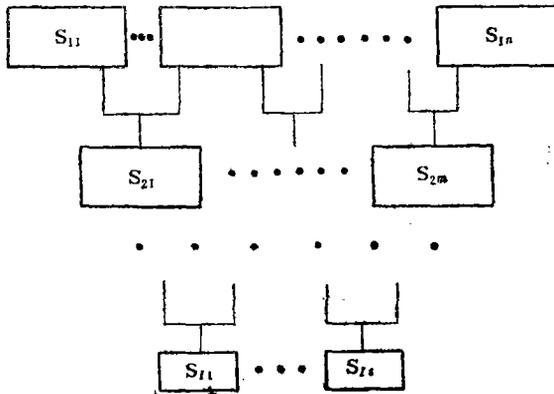
一个编译程序的高效有两重意思。一是目标程序的高效,即产生的目标指令其运行效率高。对于那些编译一次而运行多次的“一劳永逸”式的程序,目标效率的高效非常令人关注。但是,从银河亿次机近几年的运行实践看,真正“一劳永逸”的程序是极少的。只有一些库程序是这样。为了求得更高的效率,人们往往用汇编语言而不是用高级语言编写库程序。编译程序高效的另一重意思是编译程序本身的高效,即编译一道源程序所花费的时间要尽可能地短。末端用户的程序往往是经常调整的程序,对于他们,编译程序的高效同样十分重要。几年来银河亿次机运行的实践表明,用户对此是很关心的。如果一道源程序在亿次机上的编译不比中档机器上的编译快得多,尽管目标程序运行

速度高，用户往往也是很很不满意的。对系统而言，这也是资源的一种浪费。

用本文介绍的方法产生的编译程序GFT，其编译效率和目标效率都高，编译效率的提高尤为显著。与同类的编译程序比较，通过许多实例的统计，其目标程序运行的效率提高50%左右，而编译效率提高10倍以上。能够用软件的办法提高如此高的效率，参加其鉴定的专家们认为是非常突出的成效。

### 2. 主要编译技术

1. 逐步求精的结构：FORTRAN 77近60个语句<sup>[2]</sup>，其外观形式均各具特色。一般地，各应有一处理程序；事实也确是这样。但有些语句，只有很少的差别，稍加处理后即为同一形式。有些语句，虽然外观差别很大，但做了一定的处理后，即可转化为另一语句的形式。如读/写语句中的隐DO循环，在处理到隐DO部分时，则与DO语句的处理无异。该编译程序的代码行数，仅为同类程序的1/3~1/4，主要是这种逐步求精的工作做得精细。其逐步求精的方法可以图示如下：



其中  $l < \dots < m < n, 2 \leq l \leq 4$

实际上  $l$  仅为 5。

这种精细，不只是深思熟虑的结果，还是长期反复实践的结果，是写许多类似的编译程序的长时期内的一种逐步求精的结果。

2. 尽可能多地运用向量运算：向量机的高效其最关键的因素就在向量运算<sup>[1]</sup>，这是向量机系统程序设计者的常识。但可惜人们往往只关心目标程序一级的向量运算，而对编译程序本身中的向量运算，则认为可能性太小——编译程序的操作多为逻辑操作、位运算，向量运算难以进行。殊不知，编译程序中的向量运算也是大有潜力的。

a. 编译程序的结构设计，要做到能开发其中的向量运算。在很大程度上可以说，编译程序是表格处理程序。因此表格如何设置，各个具体表格的形式如何设计，对开发向量运算关系甚大。如：一个符号的特性往往是连续发见的，使用时又往往要用其多

数。如果把这些特性分散在不同的表格中,则无法向量运算;如果连续地有规则地排在某一表格中,则有向量运算的可能性。

b. 表格查填要尽量采用向量算法。表格查填约占编译程序运行时间的70%左右。因此表格查填的向量化对编译程序的效率影响极大。实验表明,将常用主要表格的查填由标量操作改为向量操作,其编译效率可提高1~3倍。

表格查填所涉及的操作有向量存取,向量移位,向量减,产生向量位串等。在具体应用中,应当尽量选用运算拍数少的操作。如方法之一是:

- ① 向量取,取出表中元素
- ② 左移或右移,移掉不是搜索目标的位
- ③ 从该向量中减去要搜索的目标
- ④ 产生向量位串
- ⑤ 如果未能发现匹配,则反复循环直至该表为空。

3. 表格管理:表格管理是编译程序优质高效的决定性因素。为此,有如下几条特别值得一提:

a. 表格的设置:一般地,为使编译程序高效,要进行许多的优化;为此,需要设置的表格也就越多。以前,我们作的441B-ⅡFORTRAN编译程序是一遍扫描的,与我们的151-FORTRAN编译程序比,表格要少得多,从而它的编译程序的代码行数要短好几倍,但后者的优化程度比前者高得多。一个一遍扫描的编译程序将源程序扫描一遍就产生半目标代码,当然省去了许多功夫。

如何做到效果优而表格又不庞复,这是一种工程艺术。大致说,在表格设置时,有如下几点值得注意:

- ① 一表尽可能多用
  - ② 一种性质尽可能只在一处地方反映
  - ③ 有关同一个符号(常数、变量、数组等)的特性尽可能有规则连续陈列
- 能否做到这几点,在相当的程度上取决于经验。

b. 表格的动态管理:一道编译程序所面向的源程序是成千上万的。这些源程序可以只有几行,几十行,也可以有成千上万行。在编译它们时,所需要的编译表格的大小也就很不一样,可能只要几百个机器字,也可能要几十万、上百万个机器字。如何确定其大小和如何调整其大小,是个微妙的问题。

我们过去多采用静态的办法,即在汇编编译程序本身时,就静态地确定各个表格的大小。这种方法的优点是:无需动态改变大小、开销小,程序逻辑简单;但为了应付大程序的情况,表格一般开得较大,因而在程序较小时,造成空间上的一些浪费。本编译程序采用动态表格管理的办法,即:在起始状态表格一般不大,待到某一表格要过满时,则采用动态扩大表格乃至向OS申请空间的办法。此法与前者相反,空间利用比较合理,编译程序运行时的空间不会超大。但要求一定的编译运行开销。如何选取一个适当的初值,影响着这一开销的大小。考虑到换道开销比较大,因此不到山穷水尽,并不向系统申请空间,而是尽量在各表格的现有空间之间进行调整。由于用了向量寄存器、向量存取和适当的搜索策略,因此开销是经济的。

c. 分情况处理：大致说来，表格搜索有顺序的和非顺序的（对半、杂凑等）两种方法。前者省空间、较易插入新项，但查找费时间；后者较费空间、插入新项费事，但查找较快。本编译程序的大多数表格为顺序表格，但有分析地选取了若干查找多的表格为非顺序表格，并施以适当的搜索方法，对提高编译程序的效率大有裨益。

这里的“分情况处理”还含有另外两重意思：一是有些表格（如库程序表）是只查不填的，它们的形成是随同编译程序一道汇编而成的。二是那些在编译过程中其大小不变或变也不大的表格集中在用户工作区的一端（相对不变端），而那些其大小在编译过程中变化较大的表格则接在它们之后。

这些措施均有益于省时、高效。

4. 程序结构：较长时间以来，人们都乐于将一些大程序分成若干个适当大小（如500行左右）的块，以利程序的易读和可维。这种结构程序无疑是对的，是人们工作中的经验总结。但如果死板地将它移用到向量流水线机器的程序结构中，则并不恰当。

分块带来了如下问题：一是指令缓冲区的可用率下降，二是转子带来的开销（保护、恢复等）大大增多。实践表明，不合理的分块会使效率下降20%甚至更多。

因此，象编译程序这样的长期运行的系统程序，应当尽量少分块。这里的少分块指的是执行结构；至于书写结构，为了易读和可维，应当清晰地分块，而且要有必要的注解说明。

用这一思想来处理库程序，则应当尽量多用插入式（in-line）调用，而少用转子式调用。事实上，大多数常用库程序（如内部函数）程序行数不多，用插入式调用是方便而高效的。

还应当提倡巨型机用户多用插入式程序设计方法。这样可能耗费了较多的空间，但巨型机最本质的思想不就是以空间换取时间吗？！况且，如果程序组织合理，巧妙地使用控制语句（DO和IF），也会使空间的使用做到恰如其份。

感谢慈云桂、陈立杰的帮助及高文国、杨学军、胡子昂、陈晓桦、赵荣彩、吴薇、杨乙龙等同志的合作和辛劳。

### 参 考 文 献

- [1] 杨桃栏，FORTRAN语言中一种有效的数组处理结构，电子学报，1986年第一期
- [2] 国防科大研究所，GFT使用手册，内部

# Compiling Technique in with Vector GFT

Yang Taolan

## Abstract

Compiling techniques with Vector are described based on GFT-FORTRAN Compiler.

**KEY WORDS** Compile, Vector, Table

## 国防科技大学博士后科研流动站首批人员进站

1987年11月,北京航空学院计算机系教研室主任赵沁平博士进入国防科技大学计算机科学与技术学科博士后科研流动站,成为我国计算机科学与技术学科的第一位博士后研究人员。

年仅37岁的赵沁平博士,原是南京大学博士生导师陈家福教授的博士生,他的博士学位论文“并行推理系统NDPIS中动态运行部分KLND-ENGINE的设计与分析”及其构造的实验性系统在国内处于领先,与国外同类工作相比具有独到之处。他在计算机学报、自动化学报等刊物上发表了许多颇有价值的学术论文。目前,在著名计算机专家陈火旺教授指导下,赵博士正从事国家高技术项目,“863”研究课题归纳推理机、智能机的研究工作。

试办博士后科研站,试行博士后研究制度,是我国人才开发的一项重要措施。具体办法是挑选一些获得博士学位的人员进站从事一个阶段的研究工作,以拓宽知识面,进一步培养独立工作的能力,使之成为具有较高水平的科研、教学人员。它对在社会主义初级阶段加快培养经济建设需要的高级专门人才,加强学术交流,增强科研、教学队伍的活力,具有积极意义。

国防科技大学是国家科委首批批准建站的单位之一,建站的学科为计算机科学与技术,专业为计算机软件、计算机组织与系统结构。该站拥有著名计算机学部委员慈云桂教授、软件专家陈火旺教授、硬件专家胡守仁教授、陈福接教授。做为全军目前唯一的博士后科研流动站,国防科技大学担负着艰巨的任务。

(胡利民供稿)