关于串行程序并行化

蒋 作 高 毅
(云南民族大学 数学与计算机科学学院 云南 昆明 650031)

摘 要 并行程序设计主要有两种途径,即使用并行程序设计语言编写并行程序,或将串行程序并行化. 串行程序并行化是一种比较有效的并行程序设计的途径. 通过介绍并行技术的现状及相关分析的一些定义,给出了一个关于在串行程序中识别可并行执行语句的算法,论述了这一算法的意义.

关键词 串行程序;并行化;相关分析;数据相关
【中图分类号】TP 305 【文献标识码】A 【文章编号】1672—8513(2007)03-0274-03

The Analysis of the Parallelizing About Sequential Program

Jiang Zuo Gao Yi
(College of Mathematics and Computer Science, Yunnan Nationalities University, Kunming 650031, China)

Abstract: There are two mainly methods to design parallel program by using parallel language to design parallel program and parallelizing sequential program. It's necessary and not easy to be well up in structure of parallel program for designer. However many sequential programs already have been mature in past 40 years, so parallelizing sequential program is a better way to design parallel program. This paper analyzes an algorithm about recognition the parallel executable sentences in the serial program. Finally it indicates the significance of this algorithm.

Key words: sequential program; parallelizing; dependency analysis; data dependency

0 引言

并行编译技术是高性能计算机系统中不可或缺的一部分. 网络性能的迅速发展,给并行处理技术带来更广阔的应用和研究方向. 目前,各应用领域已经积累了大量的串行程序,而高性能机都采用了并行处理技术,并行程序设计又较为困难,从编程、调试到性能分析都要付出大量的劳动,因此用户希望系统软件能有力地支持并行程序设计,能将现在的大量串行程序自动转换成高效的并行程序. 很多应用领域中对计算能力的需求越来越高,在很长一段时间内将依靠大规模并行处理来解决. 作为将来主流计算工具的并行计算机面临的问题是缺乏运行于其上的软件支持. 并行化编译的任务是:给定一个在单处理机上运行较长的串行程序和一台具有多个处理器可同时工作的并行计算机,目的是将串行程序分解成若干个能并行执行或至少能重叠执行的代码段,使其在并行机上能更快地运行. 程序语言经历了40多年的发展,在许多领域中,串行程序已经非常成熟,把这些成熟的串行程序并行化(parallelizing),一方面可以利用这些成熟的串行程序,另一方面在时间和效率上都得到了很大的提高. 下面讨论如何识别串行程序中可并行执行语句.

1 相关性定义

冯·诺依曼体制的核心概念是存储程序的工作方式. 在串行程序相关性分析中,文中仅考虑变量.

下面先介绍一些定义:

假定j语句继i语句之后执行,其定义及符号约定如下:

①流相关:如果从i到j存在执行通路,而且如果i至少有一个输出可供j用作输入,则j语句与语句i流相关(Flow_Dependency);

②输出相关:如果i,j两语句能产生(写)同一

* 收稿日期:2007-03-16.
作者简介:蒋作(1979~),男,硕士,助教,主要研究方向:软件工程、操作系统和并行技术.

输出变量,则两者是输出相关(Out_put_Dependency);

③控制相关:如果语句j的执行依赖于语句i(i必须在j之前执行),则语句j与语句i控制相关(Control_Dependency);

④反相关:如果语句j紧接语句i,而且如果的j输出与的i输入重叠,则语句j与语句i反相关(Antidependency);

为了叙述简洁起见,如果i语句与j语句存在上面其中一种相关,则说i语句与j语句数据相关,记之为$<i,j>$.

2　判断可并行执行语句的算法

假定程序没有语法错误,并给每一个语句顺序编号.对复合语句、函数、过程,只对其编一个号,不考虑子语句.复合语句下面的子语句、函数、过程,可以再用同样的方法来分析.下面的算法都是用SPARKS语言来描述的.

下面的算法是用来判断语句的相关性的,如果i语句和j语句数据相关,我们用偏序$<i,j>$来表示.

```
算法一:判断语句 i 和语句 j 的相关性
procedure DEPENDENCY(i,j)//判断语 i,j
                  //的相关性,A 为一集合
integer i,j
if i 语句至少有一个输出可供 j 语句作输入
then A←A∪<i,j> endif
if i,j 语句输出同一变量
then A←A∪<i,j> endif
if j 语句的执行依赖于 i 语句
then A←A∪<i,j> endif
if 语句 j 输出与的语句 i 输入重叠
then A←A∪<i,j> endif
end DEPENDENCY
算法二:判断一程序内所有语句的相关性
procedure PROGRAM_DEPENDENCY
integer i,j
read(program_name) //输入程序名为 //program_name 的程序
A←Φ
i←1;j←2
while 程序不结束 do
    while i<j do
    DEPENDENCY(i,j)
    i←i+1
    repeat
    j←j+1;i←1
repeat
num←j//记录该程序的语句个数
end PROGRAM_DEPENDENCY
算法三:计算语句节点的入度
procedure GRAPH_IN
integer i,j,B(1:n)//根据语句相关性构造//B(i)记录 i 语句的入度
i←1;j←2
while j≤num do //num 是程序的语句个数
    while i<j do
    if <i,j>∈A then B(j)←B(j)+1 endif
    i←i+1
    repeat
    j←j+1;i←1
repeat
end GRAPH_IN
算法四:计算语句节点的出度
procedure GRAPH_OUT
integer i,j,b(1:n)//根据语句相关性构造//b(i)记录 i 语句的入度
i←1;j←2
while i≤num do //num 是程序的语句个数
    for j←i+1 to num
    if <i,j>∈A then b(i)←b(i)+1 endif
repeat
    i←i+1;
end GRAPH_OUT
算法五:扫描程序,识别出可并行执行的语句存于数组 C 中
procedure PARALLELIZING
integer i,j,k,num,C(1:n)//数组 C 记录分层//的情况
call PROGRAM_DEPENDENCY //调用函数
call GRAPH_IN //调用函数
call GRAPH_OUT //调用函数
j←1
while 数组 b 的值不全为 -1 do //用出度数//组全为 -1 作为结束标志
        i←1
    while i≤num do
    if B(i)=0 then C(j)←i;j←j+1 endif
    i←i+1
```

```
repeat
C(j)←0;j←j+1 //用0把层与层分开
i←1
while i≤num do
if B(i) =0 then k←i+1
                while k≤num do
                if <i,k> ∈A then B(k)←B
(k) -1 endif//i语句的入度减1
                k←k+1
                repeat
                b(i)←-1//作一标记
endif
i←i+1
repeat
repeat
print(C)
end PARALLELIZING
```

3 结语

PARALLELIZING算法把可并行执行的串行语句识别出来,有助于并行编译器的开发.参照这些可并行执行的语句,还有助于把这些串行程序转化为并行程序.但相关性测试是一个NP-完全问题.相关性测试通常采用比较保守的策略,若不能证明任何相关性是不存在的话,则就认为它们是存在的.还由于种种原因,其中最本质的是由于客观事物本身内在联系的复杂性,使的并行编译技术的研究还面临许多复杂和艰难的问题.

参考文献:

[1] 沈志宇,胡子昂,廖湘科,等.并行编译方法[M].北京:国防工业出版社,2000.
[2] 陈国良.并行计算——结构·算法·编程[M].北京:高等教育出版社,1999.
[3] WOLFE M, BANERJEE U. Data Dependence and Its Application to Parallel Processing[J]. Int J Parallel Programming, 1987, 16(2):137-178.

(责任编辑　万志琼)

关于串行程序并行化

作者： 蒋作， 高毅， Jiang Zuo， Gao Yi
作者单位： 云南民族大学数学与计算机科学学院,云南昆明,650031
刊名： 云南民族大学学报(自然科学版)
英文刊名： JOURNAL OF YUNNAN NATIONALITIES UNIVERSITY(NATURAL SCIENCES EDITION)
年，卷(期)： 2007,16(3)
被引用次数： 1次

参考文献(3条)

1. 沈志宇;胡子昂;廖湘科 并行编译方法 2000
2. 陈国良 并行计算-结构·算法·编程 1999
3. WOLFE M;BANERJEE U Data Dependence and Its Application to Parallel Processing 1987(02)

本文读者也读过(10条)

1. 姚辉萍 串行程序并行化及其在桌面网格中的应用[学位论文]2010
2. 王伟.WANG Wei 模拟退火算法的并行化策略研究[期刊论文]-电脑知识与技术2008,3(25)
3. 陈鹏飞.王长山 基于串行FastLSA的并行算法[期刊论文]-计算机工程2004,30(19)
4. 廖绍雯.曹蕾.Liao Shao-wen.Cao Lei 异种操作系统间进行软件移植的部分问题研究[期刊论文]-河西学院学报2008,24(5)
5. 王璐.梁涛.王文义.WANG Lu.LIANG Tao.WANG Wen-yi FFT算法的并行化性能分析[期刊论文]-中原工学院学报2010,21(5)
6. 江文毅 串行程序并行化的探讨[会议论文]-1999
7. 郭龙.陈闵中.叶青.GUO Long.CHEN Hong-zhong.YE Qing 构造串行程序对应的并行任务(DAG)图[期刊论文]-计算机工程与应用2007,43(1)
8. 商磊.吕全义 数值并行迭代算法中的收敛性与并行性的权衡[会议论文]-2007
9. 黄伟.赖国明.HUANG Wei.LAI Guo-ming 基于群集系统的快速排序并行化方法[期刊论文]-现代计算机（专业版）2005(5)
10. Windows与Linux动态链接库技术研究[期刊论文]-硅谷2009(19)

引证文献(1条)

1. 吴越 串行算法并行化处理的数学模型与算法描述[期刊论文]-计算机技术与发展 2012(5)