基于优化合并准则的团划分算法

张鲁峰,何连跃, 李思昆
(国防科技大学计算机学院 607 室,湖南长沙 410073)

摘 要: 图论中的团划分算法,对于高级综合中的分配有重要意义,目前高级综合中广泛采用的是 C. J. Tseng 等提出的近似算法. 文中通过分析完全点和二分点的情况,提出了两个合并准则,指导改进了目前的团划分算法. 经模拟测试表明,新算法明显提高了划分质量.

关键词: 高级综合; 团划分; 最大完全子图; 合并准则
中图分类号: TN407 **文献标识码**: A **文章编号**: 0372-2112 (2001) 08-1104-03

Clique Partitioning Algorithms Based on Improved Merging Rules

ZHANG Lu-Feng, HE Lian-Yue, LI Si-Kun
(607 School of Computer, National University Of Technology, Changsha, Hunan 410073, China)

Abstract: Clique Partitioning is an important algorithm in data path synthesis. The present partition algorithm given by C. J. Tseng is widely used in High Level Synthesis. By analyzing the instance of Complete-point and Bi-partition point, two merge rules are presented in this paper to improve the partition algorithm. The experiments show that the partition results are improved notably by using the revised algorithm.

Key words: high-level synthesis; minimum clique partition; max complete sub-graph; merge rule

1 引言

高级综合工具的出现简化了复杂集成电路的设计过程,缩短了设计周期. 高级综合理论的研究及算法的改进对于提高综合质量非常重要.

在将行为级描述转换成寄存器传输级结构描述的综合过程中,寄存器级部件资源分配(resource allocation)是一个重要步骤. 分配过程中进行部件绑定,完成变量、计算、传值等行为描述到具体的寄存器、运算器和数据传输部件的映射. 将多个行为描述元素(数据操作、变量和数据传输)分时复用在同一硬件部件上,可以有效的提高目标电路的利用率,从而减小最终电路的面积[3]. 能够复用同一部件的元素称为相容的,相容的两个条件为:一是它们可用相同类型的部件完成;二是生存期不重叠. 元素的相容可以用相容图来表示,图中结点代表需映射的描述元素,每条边表示其连接的两个结点相容. 显然,相容图是没有自圈的无向图,任意两个结点之间至多有一条边. 相容图中的完全子图(全连接子图)中的各个结点可以映射到同一个物理部件上实现.

在不考虑互连的复杂性的前提下,将相容图划分为最小数目的完全子图,就保证了分配部件数量是最少的. 此问题在图论中称为团划分算法(或称基图剖分,Clique Partition). 最优团划分并不等价于最大完全子图问题,它也是一个 NP 完全问题[6,7].

C. J. Tseng 等提出的近似团划分算法[4],是解决该问题的典型解法,并在影响广泛的高级综合工具 HAL[5]、FACET 等中得到应用. 由于此算法可以考虑部件互连的代价,所以实用性优于 Kurdahi 等提出的寄存器分配左边算法[9,10]. 而计算效率和收敛性优于遗传、退火等算法[8]. 近年来研究者以此算法为基础,还提出了低功耗的分配算法[11]. 此算法的改进对于高级综合研究很有意义.

2 现有团划分算法缺点分析

目前的团划分算法为多次循环,每步优先合并具有最多公共邻居的边的端点. 某点称为某边的公共邻居是指从该点到该边的两个端点均有边相连. 比如图 1 中 $V1$ 为 $e34$ 的公共邻居,图 1(a) 中 $e13$ 有最多的公共邻居,算法先合并 $V1$、$V3$,合并后的点称为超结点. 两个点合并时,删除这两个点及其引出的所有边,增加一个超结点,并从超结点引边到所有的公共邻居,形成一个新图 1(b). 新图只要存在边就可以进行下一次结点合并,直至划分为不相连的超结点集合.

下面通过例子来分析现有团划分算法的缺点. 对于如图 1 所示的相容图:目前算法不能保证得到最优解. 如图 1(b) 中,$e13,4$、$e2,5$、$e4,5$ 具有相同数目的公共邻居,因此可以先合并 $V4$、$V5$,这样最终会得到三个完全子图:$\{v_1, v_3\}$,$\{v_4, v_5\}$,$\{v_2\}$.

再如图 2,目前算法有可能先合并 $V2$、$V4$ 或合并 $V3$、$V5$. 这样合并,会导致最后结果具有三个完全子图. 但最优划分结

收稿日期:2000-04-14;修回日期:2001-03-09

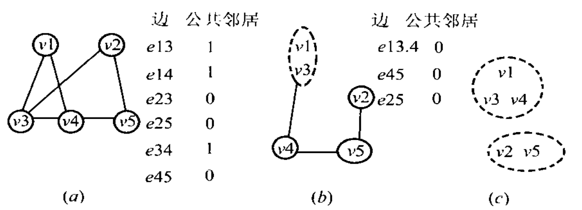


图 1　团划分算法步骤

果只有两个完全子图:{ v_1, v_2, v_3} ,{ v_4, v_5, v_6}. 为避免这两种情况,本文提出了三个合并准则,针对完全点和二分点的情况对于算法进行了改进.

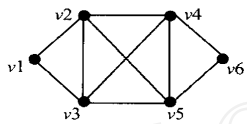


图 2

3　完全点合并准则

定义 1　如果结点合并后,得到的新图的团划分子图数目不比原图的团划分子图数目多,则称该结点合并为最佳合并.

定理 1　从相容图中删去某点及其相连的所有边,不会导致最优完全子图划分中的子图数目增加.

证明　相容图 G, a 为 G 中任意结点, G 为从 G 中去掉 a 及其相连边得到的图. 假设 G 经过最优划分后划分为 n 个完全子图 $G1$、$G2$、…、Gn,且 $G1$ 包含 a. 从 $G1$ 中去掉 a 后成为图 $G1'$,如果 $G1$ 非空,则 $G1'$、$G2$、…Gn 必为 G 的一个完全子图划分. 如果 $G1$ 为空,那么 $G2$、…、Gn 为 G 的一个完全子图划分. 因此 G 的完全子图数不会大于 G 的完全子图数.

定义 2　相容图中某点如果只有一条边与它相连,则该点称为孤立点.

定理 2　如果相容图含有孤立点,则合并孤立点与其连接点为最佳合并.

由定理 1 易证.

准则 1　相容图 G 中存在孤立点则优先合并孤立点及其邻接点.

根据这一准则,图 1(b) 进行完全子图划分时就应先合并 $V13$、$V4$ 或者合并 $V2$、$V5$,而不会先合并 $V4$、$V5$. 因此就可得到最佳划分图 1(c).

定义 3　相容图中结点 a 有 n 个邻接点,且这 n 个点构成完全子图,则称 a 点为 n —完全点.

从定义 2 中可以看出,孤立点其实是 1 —完全点.

定义 4　有两个图 $G=(V, E)$ 与 $G'=(V', E')$, G' 是 G 的子图. 对任意 $v1$, $v2$ 属于 V',如果 $v1$, $v2$ 间有边 $e12$ 属于 E,必有 $e12$ 属于 E',则称 G' 为 G 的点子图.

定理 3　有两个图 G 与 $G1$,如果 $G1$ 是 G 的点子图,则对这两个图都进行最优划分时, $G1$ 的完全子图数不会比 G 的完全子图数多.

证明大致同定理 1 的证明.

定义 5　相容图 G 中 n 个点{ $a1$, $a2$, …, an},如果{ $a1$, $a2$, …, an}两两互邻,则合并此 n 个点且删去它们引出的所有边的过程称为从 G 中划分出{ $a1$, $a2$, …, an}.

定理 4　相容图 G 中包含 n —完全点 a,则合并 a 点及其相邻接点为最佳合并.

证明:考虑 n —完全点 a,{ $a1$, $a2$, …, an}为 a 的连接点,划分中 a 只能与{ $a1$, $a2$, …, an}中的若干个点构成一个完全子图. 设最优划分中 a 与{ $a11$, $a12$, …, $a1s$}构成完全子图,从 G 中划分出此 $S+1$ 个点,记剩下的图为 $G1$. 再记从 G 中划分出{ a, $a1$, $a2$, …, an}后图为 $G2$. 则 $G2$ 为 $G1$ 的点子图,根据定理 3,最优划分后 $G2$ 的完全子图数不会比 $G1$ 的多. 所以划分出{ a, $a1$, $a2$, …, an}是最佳的.

准则 2　如果图 G 中存在 n —完全点,则先划分出该完全点及其所有相邻点.

图 2 中,点 $V1$、$V6$ 均是 2 —完全点,所以根据准则 2,应该先划分出 $V1$、$V2$、$V3$.

4　二分点合并准则

定义 6　相容图 G 中一点 a,只有两点 b、c 与其相邻,并且 b、c 不相邻. 则称 a 为 G 的二分点.

定理 5　合并相容图中的二分点及任意一个相邻点得到的划分,不会比将 a 独立划分出来差. 根据定理 1 容易证明.

因此如果图 G 存在二分点,则有理由把二分点与其中之一的相邻点作为完全子图划分出来. 但是不能保证与任意相邻点合并,都是最优合并. 如图 3 中对于二分点 $V4$,合并 $V2$ 和 $V4$ 会导致 3 个完全子图:{ $V1$, $V3$} ,{ $V2$, $V4$} ,{ $V5$} ,而实际上合并 $V4$ 和 $V5$ 只会有两个完全子图:{ $V1$, $V2$, $V3$} ,{ $V4$, $V5$}.

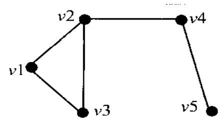


图 3

优先取度数较少的相邻点与二分点合成一个完全子图,因为这样划分出二分点,剩余的子图会有较多的边. 对于顶点数相同的图,边的总数越多,完全子图数较少的概率越大. 根据上面的思想,总结出准则 3:

准则 3　如果图 G 中包含二分点,则将二分点与其具有较少度数的相邻点作为完全子图划分出.

该准则并不一定是最优准则,但是实验结果显示,在绝大多数情况下,这种合并方式优于将二分点和度数较小的邻点合并或单独将二分点划分出来.

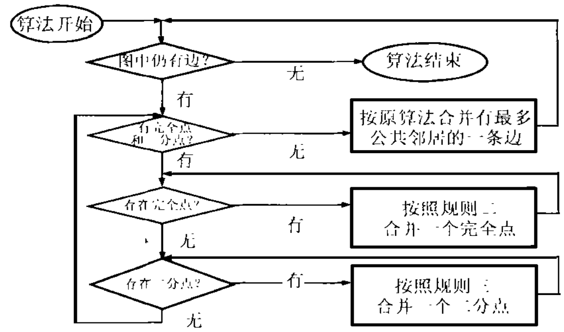


图 4　改进算法流程

5　实现及实验结果

在改进后的团划分算法中,先判断局部图中是否存在完

全点,若存在则依据准则 2 进行划分;接着判断局部图中是否存在二分点,有则用准则 3 划分;否则根据原算法进行结点合并. 因为准则 2 为最优准则,实现中准则 2 优先于准则 3 使用. 算法用 C + + 语言编码实现,程序量 1,000 行左右.

表 1 随机图结果

		平均子图数	改进个数	退步个数
1000 个平均度数 10 的相容图	原算法	43.792		
	采用规则 2	43.674	108	14
	采用规则 3	43.174	518	132
	采用两个规则	43.070	561	112
1000 个平均度数 30 的相容图	原算法	30.502		
	采用规则 2	30.347	152	32
	采用规则 3	30.091	413	94
	采用两个规则	29.898	500	83
1000 个平均度数 50 的相容图	原算法	22.640		
	采用规则 2	22.574	79	15
	采用规则 3	22.273	379	89
	采用两个规则	22.189	425	87

为了对算法进行评价,进行了相应的对比实验. 表 1 给出了采用随机生成的 3000 个 100 结点的相容图作为算法输入,原算法与使用规则 2,使用规则 3,同时使用两个规则的三种改进算法的划分结果的对比. 还选择了国外数篇有关图论算法的论文中例图作为测试输入,表 2 给出部分结果. 为了验证规则 3,我们对比了在二分点与较多度数的邻点合并和与较少度数的邻点合并的差异,表 3 给出了结果.

表 2 论文中图的结果(图的来源有 Stanford 图库,公路图,N 皇后问题)

图	原算法	改进算法
Le450_5c	149	144
Le450_15a	96	88
Le450_25c	70	64
Miles750	15	12
Anna	82	80
Queen11_11	23	22

表 3 验证规则 3 的对比实验

		平均子图数	改进个数	退步个数
1000 个平均度数 30 的相容图	二分点与较多度邻点合并的算法	32.484		
	采用规则 2	30.091	933	9

实验结果显示:

(1) 改进后的算法结果明显优于改进之前的算法. 并且得到改进的图比例较大(50 %左右),改进算法具有普遍意义.

(2) 两个规则都对结果改进有贡献. 相比来说,规则 3 的改进比规则二大.

(3) 两个规则的叠加性好,综合运用能得到比单独运用更好的结果.

(4) 在绝大多数划分中,二分点与度数较小的邻点合并是较好的选择.

6 小结

本文给出的两个准则,并以此对团划分算法进行了有效优化,经测试表明改进效果明显. 缺点是算法的复杂性较高,需要比原算法更多的计算时间. 目前工作还处于算法理论研究阶段,要将其改进为实用的电路综合工具,仍有许多工作要做.

参考文献:

[1] 向东. 并行测试的一种新策略 —测试段划分 [J]. 电子学报, 1999,2.

[2] 李丽,魏少军,杨之廉. 基于图的可测性算子资源分配算法 [J]. 清华大学学报,1999,supp.1.

[3] 袁小龙,高德远. 高级综合中寄存器合并问题的研究 [J]. 计算机工程,1999,7.

[4] C J Tseng,D P Siewiorek. Automated synthesis of data paths in digital systems [J]. IEEE Trans. ICCAD,1986,7.

[5] P G Pualin,et al. HAL:a multi-paradigm approach to automatic data path synthesis [A]. 23rd DAC 1986 [C].

[6] P Crescenzi,V Kann. A compendium of NP optiomization problems [R]. Technical report,Royal Institute of Technology,Stocholm,1998.

[7] Lars Engebretsen,et al. Towards Optimal lower bounds for clique and chromatic number [A]. 27th ICALP [C],2000.

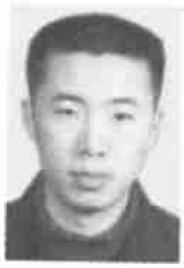
[8] Naren Narasimhan,et al. Allocation and binding for data path synthesis using a genetic algorithm approach [A]. Proc. IEEE VLSI Design1996 [C].

[9] YOUN-LONG LIN. Recent developments in high-level synthesis [J]. ACM Trans. on Design Automation of Electronic Systems. 1997,7.

[10] F J Kurdahi,et al. Real:a program for register allocation [A]. 24th DAC [C],1987.

[11] Jui-Ming Chang,et al. Register allocation and binding for low power [A]. 32ed DAC '95 [C]. 1995.

作者简介:



张鲁峰 男. 1974 年 9 月生于山东泰安,现为国防科技大学计算机学院博士研究生,主要研究方向为 EDA,嵌入式系统设计.

何连跃 男. 1971 年生,2000 年于国防科技大学计算机学院获计算机应用博士学位,主要研究方向为计算机仿真、分布虚拟环境.

李思昆 男. 1941 年生,国防科技大学计算机学院教授、博士生导师,主要研究方向为 CAD、集成电路设计、虚拟现实、分布虚拟环境.