

构造对角 Ramsey 图的 DNA 算法设计^{*}

耿修堂, 陈智华

(华中科技大学 控制科学与工程系, 武汉 430074)

摘要: Ramsey 数问题是一个著名的组合优化问题, 同时也是一个 NP 完全问题。构造对角 Ramsey 图是一个难处理的计算问题, 使用穷举的算法来构造对角 Ramsey 图必然导致计算量的指数爆炸, 穷举的 DNA 算法也不例外。提出了一个构造对角 Ramsey 图的递阶式 DNA 粘贴—剪接算法, 该算法通过逐个添加顶点的思想, 逐步删除了问题的绝大部分非解, 在一定程度上缓解了问题解的空间扩散。特别地, 专门针对对角 Ramsey 数 $R(5, 5)$ 的 43 阶 Ramsey 图的构造问题进行了计算分析, 分析结果充分地肯定了该算法的有效性。

关键词: DNA 计算; Ramsey 图; NP 完全问题; 粘贴模型; 剪接模型

中图分类号: TP183

文献标志码: A

文章编号: 1001-3695(2009)03-0827-05

Design of DNA algorithm to construct diagonal Ramsey graph

GENG Xiu-tang, CHEN Zhi-hua

(Dept. of Control Science & Engineering, Huazhong University of Science & Technology, Wuhan 430074, China)

Abstract: Ramsey number problem is a famous optimization problem, and is also a NP complete problem. So constructing diagonal Ramsey graph is an intractable computation problem. To solve the problem with exhaustion method, the space complexity will grow exponentially, and the instance is similar with primitive DNA algorithm. This paper proposed a recursive DNA sticker-splicing algorithm to construct diagonal Ramsey graph. The proposed algorithm could delete some incorrect solutions by adding the vertices stage by stage. As a result, the proposed algorithm relieved the need of space computation. Particularly, computational complexity of the presented algorithm was analyzed with the construction of Ramsey graph with 43 vertices for $R(5, 5)$. The analysis results of the algorithm show that the presented algorithm is effective.

Key words: DNA computing; Ramsey graph; NP complete problem; sticker model; splicing model

0 引言

DNA 计算是一种以 DNA 与某些相关的生物酶等作为基本材料的、基于某些生化反应原理的一种新型的分子生物计算方法。自从 Adleman^[1] 在 1994 年完成的开创实验以来, 众多学者加入到 DNA 计算研究领域, 例如 1995 年, Lipton^[2] 提出了求解可满足性问题的 DNA 算法; 1997 年, Ouyang 等人^[3] 提出了求解最大团问题的 DNA 算法; 2006 年, Li 等人^[4] 提出了求解 3-可满足性问题的 DNA 算法。

以完成对角 Ramsey 图的构造问题为目的, 本文提出了一个对角 Ramsey 图构造问题的递阶法 DNA 粘贴—剪接算法。各种 DNA 算法的第一步通常需要生成一个初始的数据池, 这个数据池包含正确、也包括不正确的解; 然后使用一系列对应的 DNA 操作来删除这些不正确的解, 从而保留了正确的解; 最后通过 DNA 检测操作把这些解检测出来, 即问题的解。然而, 这种穷举的搜索方法会受到问题规模的限制, 换句话说, 这种穷举的搜索方法会导致指数增加的 DNA 数量。正是因为存在这个瓶颈, 本文提出了一个构造对角 Ramsey 图的递阶法 DNA 粘贴—剪接算法。该算法采用逐步添加顶点、分层次分隔非解的思想, 从而在一定程度上克服了空间爆炸问题。

1 Ramsey 数问题

对角 Ramsey 数问题是一个经典的图论问题。对该问题的研究, 除了理论上的推导证明外, 众多学者也以构造 Ramsey 图方法展开了对该问题的研究。由于该问题的求解难度非常大, 要想得到新的结果或一个新的界也是一件非常困难的事情。本文以 DNA 计算为研究工具, 提出了一个构造对角 Ramsey 图的 DNA 粘贴—剪接算法。

对角 Ramsey 数问题可以用约会问题来解释, 即如何邀请最少的客人, 使得至少有 m 个人两两相互认识或至少有 m 个人两两相互不认识。图论学中是这样定义这个问题的: $R(m, m)$ 是指最小的正整数 n , 使得任意的 n 阶图, 要么含有 m 阶团, 要么含有 m 阶独立集。文献[5]给出的另外两个定义如下:

定义 1 对于一个阶为 m 的完全图 K_m , 对角 Ramsey 数 $R(K_m, K_m)$ 定义为满足下面条件的最小正整数 n , 使得: 若 K_m 的每条边都染成红色或蓝色, 无论如何染色, 必定可以找到一个红色边的子图 K_m 或一个蓝色边的子图 K_m 。这里 $m \leq n$ 。

定义 2 设 K_m 是一个阶为 m 的完全图, 对角 Ramsey 数 $R(K_m, K_m)$ 是指满足下面条件的最小正整数 n , 使得: 对任一 n 阶的图 G , G 必包含一个同构于 K_m 的子图或者 G 的补图包含一个同构于 K_m 的子图。

收稿日期: 2008-05-02; 修回日期: 2008-07-30 基金项目: 国家自然科学基金资助项目(60373089, 60674106, 60533010)

作者简介: 耿修堂(1975-), 男, 安徽人, 博士, 主要研究方向为优化算法、航空电子系统、机械设计(gxt1028@163.com); 陈智华(1976-), 女, 广西人, 讲师, 博士, 主要研究方向为智能优化、密码与信息安全、DNA 计算。

众所周知,对于 $m > 4$ 的对角 Ramsey 数精确界的确定问题是一个极其困难的问题。表 1 给出了部分经典 Ramsey 数的已知结果。由表 1 可知,对角 Ramsey 数 $R(3,3), R(4,4)$ 的精确值已经确定;对角 Ramsey 数 $R(5,5)$ 的范围在 43 与 49 之间;对角 Ramsey 数 $R(6,6)$ 的范围在 102 与 165 之间;而对角 Ramsey 数 $R(7,7)$ 的范围在 205 与 540 之间;等等。

表 1 经典 Ramsey 数的最新结果 ($k \leq 10, l \leq 10$)

$R(k, l)$	$l=3$	$l=4$	$l=5$	$l=6$	$l=7$	$l=8$	$l=9$	$l=10$
$k=3$	6 ^[6]	9 ^[6]	14 ^[6]	18 ^[7]	23 ^[8]	28 ^[9]	36 ^[10]	40 ^[11] 43 ^[12]
$k=4$		18 ^[6]	25 ^[13]	35 ^[14] 41 ^[13]	49 ^[15] 61 ^[16]	56 ^[17] 84 ^[18]	73 ^[19] 115 ^[10]	92 ^[20] 149 ^[21]
$k=5$			43 ^[22] 49 ^[13]	58 ^[14] 87 ^[23]	80 ^[24] 143 ^[25]	101 ^[20] 216 ^[21]	125 ^[17] 316 ^[25]	143 ^[17] 442 ^[16]
$k=6$				102 ^[8] 165 ^[16]	113 ^[17] 298 ^[26]	129 [*] 495 ^[26]	169 ^[17] 780 ^[27]	179 ^N 1 171 ^[26]
$k=7$					205 ^[28] 540 ^[29]	216 ^N 1 031 ^[26]	236 [*] 1 713 ^[26]	289 ^N 2 826 ^[16]
$k=8$						282 ^[30] 1 870 ^[16]	317 ^N 3 583 ^[27]	377 ^N 6 090 ^[27]
$k=9$							565 ^[32] 6 588 ^[32]	580 ^N 12 677 ^[26]
$k=10$								798 ^[31] 23 556 ^[33]

表 1 中“*”表示由 Shao 近期构造出来的两个新的下界;“N”表示对应的研究文献有待于添加。关于对角 Ramsey 数的研究,在理论上会涉及更多的复杂内容,这些理论知识超出了本文的研究范围。这里,就对角 Ramsey 图的构造问题,谈谈笔者的一些研究思路。关于对角 Ramsey 数的研究,如 $R(5,5)$,国内学者许进^[34]提出了一个猜想: $R(5,5)$ 的 Ramsey 图是个自补图,并且由 Shao 验证了 $R(5,5)$ 的 Ramsey 图不可能是正则自补图。同时,许进等人^[35]也提出了一个很有价值的问题:自补图构造和 Ramsey 数的关系问题。笔者认为,研究 Ramsey 数问题,特别是对角 Ramsey 数问题,从自补图角度来考虑这个问题很可能是一个不错的突破口。

这里,有必要对自补图的研究作一个简单阐述。关于自补图的构造问题,自从许进等人^[35]在这方面做的工作以来,日本学者^[36]在这方面做了一些补充的工作。但是,自补图构造问题的研究,一个重要的预备知识应该是图的同构判定问题。因为一个有效的图同构判定的 DNA 算法,对于构造对角 Ramsey 图 DNA 算法的设计是至关重要的。

Hsieh 等人^[37]提出了图的同构问题的一个 DNA 粘贴算法。在他们的工作中,提出了图的同构判定问题可以使用 DNA 粘贴算法有效地得到处理;特别地,他们提出的算法可以同时完成一组图的同构判断问题,这对研究 Ramsey 图的构造问题很有利。具体地说,他们主要做了两方面的工作:a) DNA 编码上他们做了大量的工作,并且结果不错;b) 在初始解的生成上,他们做了比较不错的工作,可以大大降低实验误差。总之,他们设计的 DNA 算法操作比较简便,极大地降低了实验难度,其研究工作无疑是对本文研究工作的有力支持,这也肯定了本文的研究工作。

对于对角 Ramsey 构造问题的研究,对于 $R(3,3)$ 和 $R(4,4)$ 而言,当前的电子计算机可以轻松地完成这个任务。但是,当研究 $R(5,5)$ 时,即便是使用超级的电子计算机,也是无能为力的。正是基于这个现状,本文设计了一个对角 Ramsey 图构造问题的 DNA 算法。为了详细地描述该算法,本文以 $R(3,3)$ 为研究对象,展开对角 Ramsey 图构造问题的 DNA 算法的设计。所有处理过程是并行的,且该方法适合其他对角 Ramsey 图构造问题的算法设计。

2 递阶法粘贴—剪接算法

使用 DNA 计算求解难处理的计算问题时,所面临的最大的困难就是初始数据池会随着问题的规模增加呈指数增长。例如,当研究 Ramsey 数 $R(5,5)$ 时,对 43 个点的所有图进行 Ramsey 图判定时,使用最原始的穷举算法逐个进行判定,问题的计算量为 2^{903} ! 这样的工作量是传统的电子计算机遥不可及的。

即便是使用 DNA 计算求解这样的问题,不使用改进的 DNA 算法也是无能为力的。为了降低问题的初始解空间,本文设计了对角 Ramsey 图构造问题的递阶法 DNA 粘贴—剪接算法,这种剪接操作如图 1 所示。该算法以各种存储链为处理单元,单元内部使用粘贴操作;各种存储链之间使用的是连接操作,从而达到递阶地处理问题的目的。

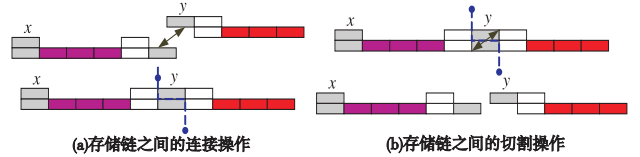


图 1 存储链之间的剪接操作描述

对角 Ramsey 图构造问题的递阶法 DNA 粘贴—剪接算法的主要过程如下:

- a) 使用足量的存储链 1 以及对应的粘贴链生成全部 m 阶的图,对这些生成链进行非解分隔,最后得到一个可行解的集合。
- b) 以 a) 所获得的解为基础,使用足量的存储链 2 以及对应的粘贴链生成全部 $m + 1$ 阶的图,对这些生成链进行非解分隔,最后得到一个新的可行解集合。
- c) 同理,以 b) 所获得的解为基础,使用足量的存储链 i 以及对应的粘贴链生成全部 $m + i - 1$ 阶的图,对这些生成链进行非解分隔,最后得到一个新的可行解集合。直到对存储链 $m - n + 1$ 进行同样的处理之后,递阶法操作全部结束。
- d) 这时候的所有可行解,即问题的最终解。具体地说,如果试管里没有生成链,表明没有发现任何 Ramsey 图;如果存在生成链,说明找到了一个 Ramsey 图。最后,可以通过检测操作检测出任意一个解。

3 构造对角 Ramsey 图的 DNA 算法

具体到对角 $R(3,3)$ 的 Ramsey 图构造问题,该算法需要 3 类存储链、10 类粘贴链。第一类存储链存储三个顶点的全部可能存在的图;第二类存储链存储第四个顶点与全部三个顶点形成的图的连接关系;第三类存储链存储第五个顶点与全部四个顶点形成的图的连接关系。10 类粘贴链对应五个顶点之间的十条边。这三类存储链以及 10 类粘贴链如图 2 所示。

由图 2 可见,存储链 x 有三位信息;存储链 y 有三位信息;存储链 z 有四位信息。其中粘贴链 $e_{i,j}$ 表示顶点 i 与顶点 j 之间存在边。到此为止,构造 $R(3,3)$ 的 Ramsey 图的递阶法 DNA

粘贴—剪接算法的简单过程如下:

- a) 将足量的存储链 x 、粘贴链 $e_{2,1}$ 、 $e_{3,1}$ 、 $e_{3,2}$ 混合在一起进行退火操作,生成全部可能的三个顶点的图,对这些生成链 x 进行非解分隔,即删除全部团为 3 和全部独立集为 3 的图。
- b) 将等量的存储链 y 加入 a) 得到的生成链 x 中,通过退火,连接操作生成了生成链 xy ;然后加入适量的粘贴链 $e_{4,1}$ 、 $e_{4,2}$ 、 $e_{4,3}$,通过退火操作生成了全部 4 阶的图。对这些生成链 xy 进行非解分隔,即删除全部含团为 3 和全部包含独立集为 3 的图。
- c) 将等量的存储链 z 加入 b) 得到的生成链 xy 中,通过退火,连接操作生成了生成链 xyz ;然后加入适量的粘贴链 $e_{5,1}$ 、 $e_{5,2}$ 、 $e_{5,3}$ 、 $e_{5,4}$,通过退火操作生成了全部 5 阶的图。对这些生成链 xyz 进行非解分隔,即删除全部含团为 3 和全部包含独立集为 3 的图。
- d) 这时候的所有解即问题的最终解。具体地说,就是试管中要是 有解,它们既不含团 3,也不含独立集 3。

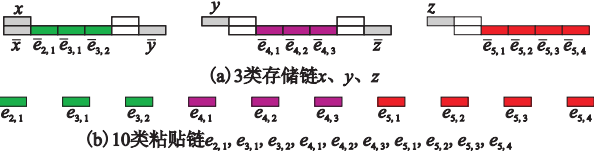


图2 3类存储链和10类粘贴链的描述

最后,该算法的详细过程如下:

- a) 试管中放入适量的存储链 x ,接着向试管中加适量的粘贴链 $e_{2,1}$ 、 $e_{3,1}$ 、 $e_{3,2}$;然后使用退火操作生成了若干生成链,这些生成链的位信息包含八种可能状态,即 $e_{2,1}e_{3,1}e_{3,2} = 000, 001, 010, 011, 100, 101, 110, 111$ 。下面通过分离操作实现非解分隔,分离工具是固定于试管内壁的核苷酸探针。这里,对应于 3 类粘贴链,总共使用了三种探针。在算法操作的第一步,所有生成链 x 放入试管 T_1 中,使用探针 $e_{2,1}$ 将试管中的生成链 x 一分为二,由试管 T_2 容纳 $e_{2,1} = 0$ 的所有生成链 x ,由试管 T_3 容纳 $e_{2,1} = 1$ 的所有生成链 x 。接着使用探针 $e_{3,1}$ 将试管 T_2 中的生成链 x 一分为二,由试管 T_4 容纳 $e_{3,1} = 0$ 的所有生成链 x ,由试管 T_5 容纳 $e_{3,1} = 1$ 的所有生成链 x ;同时使用探针 $e_{3,2}$ 将试管 T_3 中的生成链 x 一分为二,由试管 T_6 容纳 $e_{3,2} = 0$ 的所有生成链 x ,由试管 T_7 容纳 $e_{3,2} = 1$ 的所有生成链 x 。接着使用探针 $e_{3,2}$ 将试管 T_4 中的生成链 x 一分为二,由试管 T_8 容纳 $e_{3,2} = 0$ 的所有生成链 x ,由试管 T_9 容纳 $e_{3,2} = 1$ 的所有生成链 x ;同时使用探针 $e_{3,2}$ 将试管 T_5 中的生成链 x 一分为二,由试管 T_{10} 容纳 $e_{3,2} = 0$ 的所有生成链 x ,由试管 T_{11} 容纳 $e_{3,2} = 1$ 的所有生成链 x 。使用探针 $e_{3,2}$ 将试管 T_6 中的生成链 x 一分为二,由试管 T_{12} 容纳 $e_{3,2} = 0$ 的所有生成链 x ,由试管 T_{13} 容纳 $e_{3,2} = 1$ 的所有生成链 x 。使用探针 $e_{3,2}$ 将试管 T_7 中的生成链 x 一分为二,由试管 T_{14} 容纳 $e_{3,2} = 0$ 的所有生成链 x ,由试管 T_{15} 容纳 $e_{3,2} = 1$ 的所有生成链 x 。最后将试管 T_9 、 T_{10} 、 T_{11} 、 T_{12} 、 T_{13} 、 T_{14} 中的所有存储链放入清空的试管 T_1 里,其他的试管中全部是非解。本步操作的分离过程如图 3 所示。
- b) 首先加入适量的存储链 y 到算法第一步所得到的生成链 x 中;然后通过连接酶连接操作来连接生成链 x 与存储链 y 得到生成链 xy ;接着,向生成链 xy 中加入适量的粘贴链 $e_{4,1}$ 、 $e_{4,2}$ 、 $e_{4,3}$,通过退火操作生成新的生成链。此时,新的生成链 xy 共有 48 种状态,即 $e_{2,1}e_{3,1}e_{3,2}e_{4,1}e_{4,2}e_{4,3} = 001000, 001001,$

001010, 001011, 001100, 001101, 001110, 001111, 010000, 010001, 010010, 010011, 010100, 010101, 010110, 010111, 011000, 011001, 011010, 011011, 011100, 011101, 011110, 011111, 100000, 100001, 100010, 100011, 100100, 100101, 100110, 100111, 101000, 101001, 101010, 101011, 101100, 101101, 101110, 101111, 110000, 110001, 110010, 110011, 110100, 110101, 110110, 110111。分离操作使用到的探针有: $e_{2,1}$ 、 $e_{3,1}$ 、 $e_{3,2}$ 、 $e_{4,1}$ 、 $e_{4,2}$ 、 $e_{4,3}$ 。需要处理的对象为 3 类子图: $v_4v_2v_1$, $v_4v_3v_1$, $v_4v_3v_2$ 。删除子图 3 的所有图和子图 3 的所有图。对每类子图的非解删除处理过程完全类似算法第一步操作,即计算量是第一步操作的三倍。最后,试管里剩下的新解为 $e_{2,1}e_{3,1}e_{3,2}e_{4,1}e_{4,2}e_{4,3} = 001100, 001101, 001110, 010010, 010011, 010110, 011010, 011100, 011110, 100001, 100011, 100101, 101001, 101100, 101101, 110001, 110010, 110011$ 。这 18 个 4 阶图既不含团 3 的子图,也不含独立集 3 的子图。

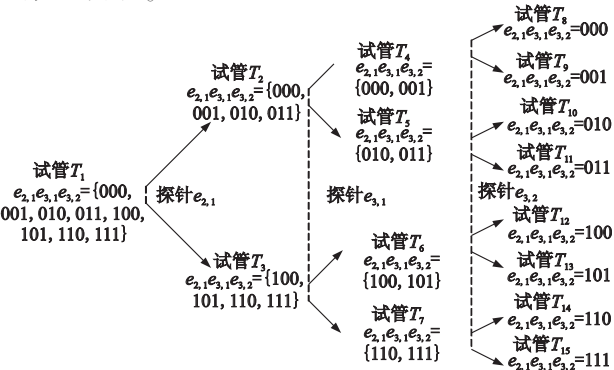


图3 分离流程的描述

表 2 所有五个顶点的图的 288 种可行解状态

生成链 xy 位信息	生成链 xyz 位信息 ($e_{5,1}, e_{5,2}, e_{5,3}, e_{5,4}$)
001100	0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1010, 1011, 1100, 1101, 1110, 1111
001101	0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1010, 1011, 1100, 1101, 1110, 1111
001110	0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1010, 1011, 1100, 1101, 1110, 1111
010010	0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1010, 1011, 1100, 1101, 1110, 1111
010011	0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1010, 1011, 1100, 1101, 1110, 1111
010110	0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1010, 1011, 1100, 1101, 1110, 1111
011010	0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1010, 1011, 1100, 1101, 1110, 1111
011100	0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1010, 1011, 1100, 1101, 1110, 1111
011110	0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1010, 1011, 1100, 1101, 1110, 1111
100001	0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1010, 1011, 1100, 1101, 1110, 1111
100011	0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1010, 1011, 1100, 1101, 1110, 1111
100101	0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1010, 1011, 1100, 1101, 1110, 1111
101001	0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1010, 1011, 1100, 1101, 1110, 1111
101100	0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1010, 1011, 1100, 1101, 1110, 1111
101101	0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1010, 1011, 1100, 1101, 1110, 1111
110001	0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1010, 1011, 1100, 1101, 1110, 1111
110010	0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1010, 1011, 1100, 1101, 1110, 1111
110011	0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1010, 1011, 1100, 1101, 1110, 1111

c) 首先加入适量的存储链 z 到算法 b) 所得到的生成链 xy 中; 然后通过连接酶连接操作来连接生成链 z 与存储链 xy , 得到生成链 xyz ; 接着, 向生成链 xyz 里加入适量的粘贴链 $e_{5,1}$ 、 $e_{5,2}$ 、 $e_{5,3}$ 、 $e_{5,4}$, 通过退火操作生成新的生成链 xyz 。此时, 新的生成链 xyz 共有 288 种状态, 如表 2 所示。步骤 c) 分离操作使用到的探针有 $e_{2,1}$ 、 $e_{3,1}$ 、 $e_{3,2}$ 、 $e_{4,1}$ 、 $e_{4,2}$ 、 $e_{4,3}$ 、 $e_{5,1}$ 、 $e_{5,2}$ 、 $e_{5,3}$ 、 $e_{5,4}$ 。需要处理的对象为 6 类子图: $v_5 v_2 v_1$ 、 $v_5 v_3 v_1$ 、 $v_5 v_3 v_2$ 、 $v_5 v_4 v_1$ 、 $v_5 v_4 v_2$ 、 $v_5 v_4 v_3$ 。这里同样删除子图 3 的所有图和子图 3 的所有图。对每类子图的非解删除处理过程完全类似前面的步骤 a), 但计算量是操作 a) 的六倍。最后, 试管中剩下的新解为 $e_{2,1} e_{3,1} e_{3,2} e_{4,1} e_{4,2} e_{4,3} e_{5,1} e_{5,2} e_{5,3} e_{5,4} = 0011011100, 0011101010, 0100111100, 0101100110, 0110101001, 0111000101, 1000111010, 1001010110, 1010011001, 1011000011, 1100010101, 1100100011$ 。这 12 个 5 阶图既不含子图 3 的子图, 也不含有独立集 3 的子图, 它们都是对角 Ramsey 数的 Ramsey 图。对应这 12 个图的 10 位信息, 对应的 12 个 Ramsey 图如图 4 所示。

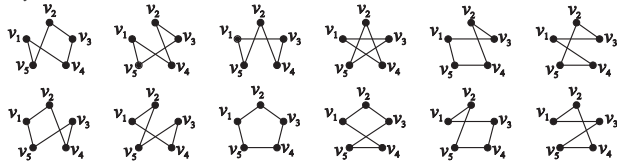


图 4 $R(3,3)$ 的 12 个 Ramsey 图

d) 步骤 c) 中已经删除了全部非解, 这时候试管中只要有生成链 xyz 就表明存在 Ramsey 图。那么, 如何读出一个生成链 xyz 的位信息呢? 该步骤就是要完成这个任务。检测最后解的 DNA 操作使用的是分离操作。这里以生成链 xyz 的值是 $e_{2,1} e_{3,1} e_{3,2} e_{4,1} e_{4,2} e_{4,3} e_{5,1} e_{5,2} e_{5,3} e_{5,4} = 0011011100$ 为例来讨论具体的解检测过程。(a) 使用探针 $e_{2,1}$ 分离生成链 xyz 为两部分, 选择有解的部分 (这里 $e_{2,1} = 0$); (b) 对有解的试管使用探针 $e_{3,1}$ 分离为两部分, 选择有解的部分 (这里 $e_{3,1} = 0$); (c) 对有解的试管使用探针 $e_{3,2}$ 分离为两部分, 选择有解的部分 (这里 $e_{3,2} = 1$); (d) 对有解的试管使用探针 $e_{4,1}$ 分离为两部分, 选择有解的部分 (这里 $e_{4,1} = 1$); (e) 对有解的试管使用探针 $e_{4,2}$ 分离为两部分, 选择有解的部分 (这里 $e_{4,2} = 0$); (f) 对有解的试管使用探针 $e_{4,3}$ 来分离为两部分, 选择有解的部分 (这里 $e_{4,3} = 1$); (g) 对有解的试管使用探针 $e_{5,1}$ 分离为两部分, 选择有解的部分 (这里 $e_{5,1} = 1$); (h) 对有解的试管使用探针 $e_{5,2}$ 分离为两部分, 选择有解的部分 (这里 $e_{5,2} = 1$)。 (i) 有解的试管使用探针 $e_{5,3}$ 分离为两部分, 选择有解的部分 (这里 $e_{5,3} = 0$); (j) 有解的试管使用探针 $e_{5,4}$ 分离为两部分, 选择有解的部分 (这里 $e_{5,4} = 0$)。检测结果表明这个解为 0011011100。整个检测操作如图 5 所示, 图中实线表示该部分为非空集合, 即存在生成链 xyz 。

4 算法分析

结合第 3 章 $R(3,3)$ 的 5 阶对角 Ramsey 图构造问题的详细算法描述, 分析 $R(m, m)$ 的 n 阶对角 Ramsey 图构造问题的算法复杂性。当问题规模为 n , 对角 Ramsey 数为 $R(m, m)$ 时, 该算法需要存储链种类为 S_{cl} (cl 为存储链), 则

$$S_{cl} = n - m + 1 \quad (1)$$

需要粘贴链种类为 S_{zl} (zl 为粘贴链), 则

$$S_{zl} = \sum_{i=1}^{i=n-1} i = n(n-1)/2 \quad (2)$$

需要探针种类为 S_{tz} (tz 为探针), 则

$$S_{tz} = \sum_{i=1}^{i=n-1} i = n(n-1)/2 \quad (3)$$

需要测试试管总和为 S_{sg} (sg 为试管), 则

$$S_{sg} = (n - m + 1) + \sum_{i=1}^{i=n-1} i + (2m + 1) = (n^2 + n + 2m + 4)/2 \quad (4)$$

所以, 该算法的空间种类总和为 S_{kj} (kj 为空间), 则

$$S_{kj} = S_{cl} + S_{zl} + S_{tz} + S_{sg} = (3n^2 + n + 6)/2 \quad (5)$$

由式 (5) 可知, 该算法的占用空间复杂度为 $O(n^2)$ 。该算法的分离 (Π) 操作次数为

$$S_{\Pi} = (2m - 1) (1 + C_{m-1}^{m-1} + C_{m+1}^{m-1} + C_{m+2}^{m-1} + \dots + C_{n-1}^{m-1}) + \sum_{i=1}^{i=n-1} i \quad (6)$$

从操作复杂度的角度考虑, 分离操作是该算法的主要操作。为了方便该算法复杂性的分析, 这里以分离操作复杂性来表示本文所提出算法的计算复杂性。随着问题规模 n 的增大, m 会越来越小, 即 $m < n$ 。故该算法的计算时间复杂度为 $O(mn^{m-1})$ 。

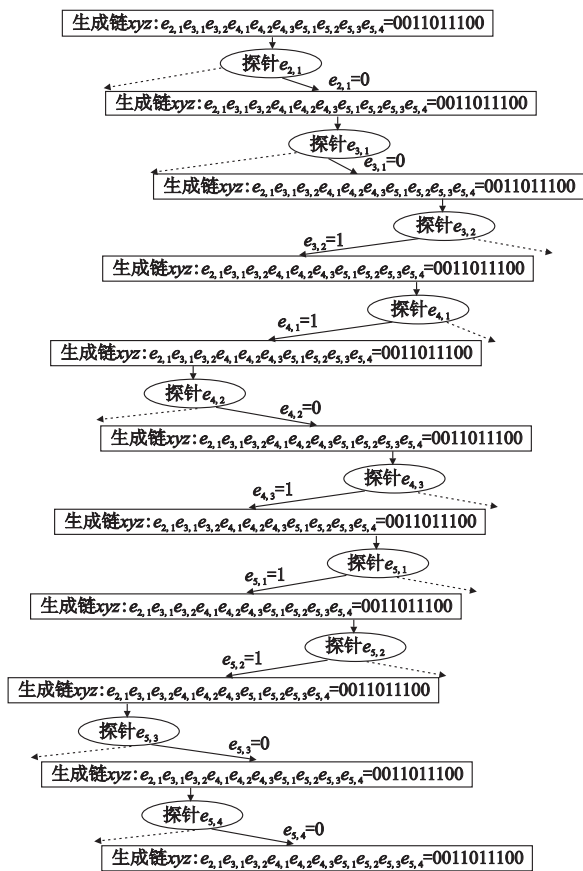


图 5 解的检测操作流程示意图

5 结束语

本文所设计的 DNA 算法利用了 DNA 计算的并行处理能力, 从而使算法的计算时间复杂度得到了极大的降低。但是计算时间复杂度为 $O(mn^{m-1})$ 的计算量也是很可怕的。特别是该算法还没有对到底需要多少存储链、粘贴链进行定量分析, 因为要弄清这个问题, 还需要大量的图论知识背景, 也是一个非常复杂的问题。好在这个问题并不影响本文的研究主题。

本文的主要工作是设计了一个对角 Ramsey 图构造问题的递阶法 DNA 粘贴—剪接算法。该算法用于构造高阶的对角 Ramsey 图时确实是很困难, 或不能实现。但是当把问题转移

到 $R(5,5)$ 这个特定问题时,的确具有令人眼前一亮的感觉。由表 1 可知 $43 \leq R(5,5) \leq 49$ 。当考虑 $R(5,5)$ 的 43 阶 Ramsey 图的构造问题时,根据式(6)可知,该算法需要的操作步数为 $8\ 663\ 382$ 。如前所述, 2^{903} 的计算空间复杂度是一个天文数据,本文使用了递阶法的处理思想,目的就是要采用逐步添加顶点、分层次地分隔非解,最终降低了这个巨大的计算空间复杂度。

由图 4 可知,在对角 Ramsey 数 $R(3,3)$ 的 5 阶 Ramsey 图的构造算法的步骤 c) 中,虽然问题规模由原来的 2^{10} 种可能降低到只有 12 种可能。但是,这 12 个图是相互同构的,也就是说,如果能够删除这些相互同构的图,该问题的规模可能将由原来的 1 024 种降低为仅 1 种可能。所以如何递阶地排除同构的图,也就是进一步降低算法的计算空间复杂度,也是本文所提出算法即将研究的重点工作。当然,该项工作也必须是基于 DNA 算法的。

根据本文综述部分对图的同构判别的 DNA 算法的分析,要想把 Hsieh 等人的思想应用于对角 Ramsey 图构造问题的 DNA 算法设计上,这不是一件仅仅把别人的算法拿过来的简单的事,还有很多的具体工作需要完成,很多困难需要克服。

参考文献:

- [1] ADLEMAN L M. Molecular computation of solutions to combinatorial problems[J]. *Science*, 1994, 266(11): 1021-1024.
- [2] LIPTON R J. DNA solution of hard computational problems[J]. *Science*, 1996, 268(28): 542-545.
- [3] OUYANG Qi, KAPLAN P D, LIU Shu-mao, et al. DNA solution of the maximal clique problem[J]. *Science*, 1997, 278(17): 446-449.
- [4] LI Da-fa, LI Xiang-rong, HUANG Hong-tao, et al. Scalability of the surface based DNA algorithm for 3-SAT[J]. *BioSystems*, 2006, 85(2): 95-98.
- [5] CHARTRAND G, ZHANG Ping. 图论导引[M]. 范益政, 汪毅, 龚世才, 等译. 北京: 人民邮电出版社, 2007.
- [6] GREENWOOD R E, GLEASON A M. Combinatorial relations and chromatic graphs[J]. *Canadian Journal of Mathematics*, 1955(7): 1-7.
- [7] GRAVER J E, YACKEL J. Some graph theoretic results associated with Ramsey's theorem[J]. *Journal of Combinatorial Theory*, 1968(4): 125-175.
- [8] KALBFLEISCH J G. Chromatic graphs and Ramsey's theorem[D]. Waterloo, Canada: University of Waterloo, 1966.
- [9] McKAY B D, MIN Zhang-ke. The value of the Ramsey number $R(3, 8)$ [J]. *Journal of Graph Theory*, 1992, 16(1): 99-105.
- [10] GRINSTEAD C M, ROBERTS S M. On the Ramsey numbers $R(3, 8)$ and $R(3, 9)$ [J]. *Journal of Combinatorial Theory*, 1982, 33: 27-51.
- [11] EXOO G. On two classical Ramsey numbers of the form $R(3, n)$ [J]. *SIAM Journal of Discrete Mathematics*, 1989, 2(4): 488-490.
- [12] RADZISZOWSKI S, KREHER D L. Upper bounds for some Ramsey numbers $R(3, k)$ [J]. *Journal of Combinat Math Combin Comput*, 1988, 12(4): 207-212.
- [13] McKAY B D, RADZISZOWSKI S P. $R(4, 5) = 25$ [J]. *Journal of Graph Theory*, 1995, 19(5): 309-322.
- [14] EXOO G. Announcement: on the Ramsey numbers $R(4, 6)$, $R(5, 6)$ and $R(3, 12)$ [J]. *Ars Combinatoria*, 1993, 35: 85.
- [15] EXOO G. Applying optimization algorithm to Ramsey problems[C]//ALAVI Y. Graph theory, combinatorics, algorithms and applications. Philadelphia: SIAM, 1989: 175-179.
- [16] MACKEY J. Combinatorial remedies[D]. Hawaii: University of Hawaii, 1994.
- [17] EXOO G. Some new Ramsey colorings[J]. *Electronic Journal of Combinatorics*, 1998, 5(1): 1-5.
- [18] EXOO G. Some applications of PQ-groups in graph theory[J]. *Discussiones Math Graply Theory*, 2004, 24(1): 109-114.
- [19] RADZISZOWSKI S, KREHER D L. Search algorithm for Ramsey graphs by union of group orbits[J]. *Journal of Graph Theory*, 1988, 12(1): 59-72.
- [20] PIWAKOWSKI K. Applying tabu search to determine new Ramsey numbers[J]. *Electronic Journal of Combinatorics*, 1996, 3(1): 1-4.
- [21] HARBORTH H, KRAUSE S. Ramsey numbers for circulant colorings[J]. *Congressus Numerantium*, 2003, 161: 139-150.
- [22] EXOO G. A lower bound for $R(5, 5)$ [J]. *Journal of Graph Theory*, 1989, 13(1): 97-98.
- [23] SPENCER T. Upper bounds for Ramsey numbers via linear programming[EB/OL]. (1993) [2007]. <http://www.mathworld.wolfram.com/RamseyNumber.html>.
- [24] CALKIN N J, ERDÖS P, TOVEY C A. New Ramsey bounds from cyclic graphs of prime order[J]. *SIAM Journal on Discrete Mathematics*, 1997, 10(3): 381-387.
- [25] HAANPÄÄ H. A lower bound for a Ramsey number[J]. *Congressus Numerantium*, 2000, 144: 189-191.
- [26] XU Xiao-dong, XIE Zheng, RADZISZOWSKI S P. A constructive approach for the lower bounds on the Ramsey numbers $R(s, t)$ [J]. *Journal of Graph Theory*, 2004, 47(3): 231-239.
- [27] XU Xiao-dong, XU Zheng, EXOO G, et al. Constructive lower bounds on classical multicolor Ramsey numbers[J]. *Electronic Journal of Combinatorics*, 2004, 11(1): 35-58.
- [28] HILL R, IRVING R W. On group partitions associated with lower bounds for symmetric Ramsey numbers[J]. *European Journal of Combinatorics*, 1982(3): 35-50.
- [29] GIRAUD G. Une minoration du nombre de quadrangles unicolores et son application a la majoration des nombres de Ramsey binaires bicolores[J]. *C R Acad Sci Paris A*, 1973, 276: 1173-1175.
- [30] BURLING J P, REYNER S W. Some lower bounds of the Ramsey numbers $n(k, k)$ [J]. *Combinatorial Theory*, 1972, 13: 168-169.
- [31] SHEARER J B. Lower bounds for small diagonal Ramsey numbers[J]. *Journal of Combinatorial Theory*, 1986, 42(2): 302-304.
- [32] SHI Ling-sheng, ZHANG Ke-min. An upper bound formula for Ramsey numbers[M]. [S. l.]: Preprint, 2001.
- [33] SHI Ling-sheng. Upper bounds for Ramsey numbers[M]. [S. l.]: Preprint, 2002.
- [34] 许进. 自补图理论及应用[M]. 西安: 西安电子科技大学出版社, 1999.
- [35] XU Jin, WONG C K. Self-complementary graphs and Ramsey numbers part I: the decomposition and construction of self-complementary graphs[J]. *Discrete Mathematics*, 2000, 223(1-3): 309-326.
- [36] KAWARABAYASHI K, NAKAMOTO A, ODA Y, et al. On separable self-complementary graphs[J]. *Discrete Mathematics*, 2002, 257(1): 165-168.
- [37] HSIEH S Y, CHEN M Y. A DNA-based solution to the graph isomorphism problem using Adleman-Lipton model with stickers[J]. *Applied Mathematics and Computation*, 2008, 197(2): 672-686.