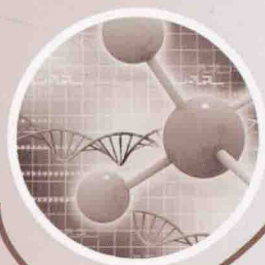


大学计算机学科学术研究进展系列丛书

DNA 计算

核酸编码原理及方法

张 凯 著



科学出版社

大学计算机学科学术研究进展系列丛书

DNA 计算核酸编码原理及方法

张 凯 著



科学出版社

北京

版权所有,侵权必究

举报电话:010-64030229;010-64034315;13501151303

内 容 简 介

DNA 计算是借助于分子生物技术进行计算的新方法。凭借着极大的存储密度和高度并行性,DNA 计算为求解 NP 完全问题等复杂组合优化问题提供了一条全新的途径,开创了广阔的应用前景。本书主要介绍 DNA 计算核酸编码原理及方法,具体包括 DNA 计算的研究进展和背景、DNA 计算的生物化学基础、DNA 编码问题及其复杂性分析、DNA 二级结构预测和最小自由能模型、隐枚举核酸序列编码算法、DNA 编码在图着色 DNA 计算中的应用。

本书可作为从事 DNA 计算、纳米结构设计、分子自组装、高性能的计算的研究学者和学生的参考书,对相关专业的教师和科研人员进行科学研究也具有一定的参考价值。

图书在版编目(CIP)数据

DNA 计算核酸编码原理及方法/张凯著. —北京:科学出版社,2014.12

(大学计算机学科学术研究进展系列丛书)

ISBN 978-7-03-042928-5

I. ①D… II. ①张… III. ①脱氧核糖核酸—计算方法—编码—研究
IV. ①O157.4

中国版本图书馆 CIP 数据核字(2015)第 000433 号

责任编辑:张颖兵 杜 权/责任校对:肖 婷

责任印制:赵 博 / 封面设计:苏 波

科 学 出 版 社 出 版

北京东黄城根北街 16 号

邮政编码:100717

<http://www.sciencep.com>

双青印刷厂印刷

科学出版社发行 各地新华书店经销

*

开本: B5(720×1000)

2015 年 3 月第 一 版 印张:11 3/4

2015 年 3 月第一次印刷 字数:210 000

定价:68.00 元

(如有印装质量问题,我社负责调换)

前 言

DNA 计算是一种基于 DNA 分子的并行计算模型,是在计算科学和分子生物学的基础上发展起来的一个新颖而极具发展潜力的交叉学科,凭借着极大的存储密度和高度并行性,DNA 计算为求解 NP 完全问题等复杂组合优化问题,提供了一条全新的途径,开创了广阔的应用前景。

近年来,随着 DNA 计算研究的不断深入,DNA 计算逐步向自组装方向发展,DNA 自组装技术已成为生物计算和 DNA 纳米结构设计的一种重要的模型。在不需要任何外部条件干预下,这些经过设计的 DNA 分子能精确地形成具有较大规模且复杂的结构,进一步完成特定问题的计算。目前在 Seeman、Adleman、Winfree、北京大学许进、上海交通大学贺林等国内外学者的研究成果的持续推动下,DNA 计算技术、DNA 自组装技术、DNA 纳米结构设计等已经从理论模型研究阶段全面发展到生物实验的实际应用阶段。

目前,DNA 计算需要解决的最基本问题是提高 DNA 计算的可靠性、有效性和可扩充性。DNA 编码质量的优劣直接决定 DNA 计算效率的高低,DNA 编码数量的多少决定了 DNA 计算可求解问题的规模,为了提高 DNA 计算的可靠性、有效性,必须设计高质量的 DNA 编码,因此 DNA 编码是 DNA 计算研究中的重要课题。

由于 DNA 计算涉及计算机、生物、数学等多个学科,所以国内对该领域深入讨论的专著很少。此外,该领域的科研成果通常发表在英文的学术期刊上,都是介绍最新的前沿技术和应用,很少涉及基础概念等内容,不适合初学者学习。著者于 2005 年起在华中科技大学攻读博士学位,研究 DNA 计算方向,2008—2010 年在北京大学信息学院博士后流动站工作,先后主持和参与了多项国家自然科学基金、国家 863 项目的研究工作,积累了一定的研究经验。撰写本书的目的是为希望了解、学习这个领域的学生和学者提供一部通俗易懂、由浅入深的学习书籍,同时也是个人科研工作的阶段总结。

本书共分为 8 章,主要内容如下。

第 1 章介绍了 DNA 计算的研究背景和最新进展,分析了目前 DNA 计算的难点问题和 DNA 计算核酸编码的作用与研究意义。

第 2 章介绍了 DNA 的分子组成和结构、DNA 计算生物化学操作和 DNA 计算的实现途径。

第 3 章对 DNA 编码约束进行规范分类,研究了编码质量、编码数量、序列长度与 DNA 计算可靠性、有效性、可扩充性之间的关系,对 DNA 编码问题的复杂性进行了分析。

第 4 章介绍了 DNA 分子的各种二级结构和 Nearest-Neighbor 热力学参数,以及最小自由能模型;提出一种动态规划回溯树搜索策略,改进了传统动态规划二级结构预测算法,有效降低了算法时间复杂度,如果满足最小自由能的二级结构不唯一,可以找到自由能相等的所有结构。

第 5 章针对 DNA 伪结预测这一难题,提出了一种基于遗传算法的平面伪结搜索方法。该算法利用连续碱基堆栈选取优化结构,采用最小自由能作为适应度函数。设计了交叉和变异算子产生新的种群个体。通过和 RNAstructure 等经典预测软件进行对比,证明了本书算法的有效性。

第 6 章提出了隐枚举核酸序列设计算法,通过将编码约束转换为整数线性规划的条件不等式对候选 DNA 序列进行约束,巧妙地产生并评价候选 DNA 序列。算法通过剪枝策略高效地搜索 4^n 的 DNA 编码解空间,找出满足约束条件的最大的 DNA 序列集合。通过和其他经典算法的比较,证明了算法的有效性和可靠性。

第 7 章研究了基于 DNA 计算的图顶点着色方法,通过对图中的顶点和颜色进行恰当的编码,设计了 129 条 DNA 序列,通过分子生物学的实验方法,成功求解了 61 个顶点的图的 3 着色问题实例。

第 8 章介绍了经典 ST-编号平面嵌入算法,提出了一种考虑 DNA 的环状和碱基堆栈等结构特点的 DNA 二级结构平面嵌入算法,对 DNA 分子进行布局调整和平面嵌入。通过算法将 DNA 二级结构自动地绘制为平面图,方便研究人员对 DNA 分子进行分析和研究。

限于著者的水平,书中的不足之处在所难免,恳请广大读者和专家指正。

本书的出版得到了国家自然科学基金(6110055, 61472293)的经费资助,在此表示感谢。在本书的撰写和出版过程中还得到科学出版社杜权编辑和其他同志的支持和帮助,在此一并表示衷心的感谢。

张 凯

2014年9月1日

目 录

第 1 章 DNA 计算的产生与发展	1
1.1 国内外研究进展	5
1.2 DNA 计算问题难点	15
1.3 DNA 计算编码方法的研究意义	17
第 2 章 DNA 计算的生物化学基础	19
2.1 DNA 的分子组成和结构	19
2.2 DNA 计算生物化学操作	23
2.2.1 DNA 的变性、复性和杂交	23
2.2.2 PCR 扩增	25
2.2.3 凝胶电泳分离	26
2.2.4 DNA 链的外切	27
2.2.5 DNA 链的内切	28
2.2.6 DNA 链的连接	29
2.2.7 特定 DNA 分子的提取	31
2.2.8 DNA 序列的测定	31
2.3 DNA 分子计算的实现途径	32
2.3.1 基于溶液反应的 DNA 分子计算	32
2.3.2 基于表面的 DNA 计算	33
2.3.3 基于 DNA 芯片的 DNA 计算	33
第 3 章 DNA 编码问题及其复杂性研究	35
3.1 DNA 编码研究的重要意义	35
3.2 DNA 编码问题的复杂性分析	37
3.3 DNA 编码约束及分析	40
3.3.1 基于汉明距离的编码约束	40
3.3.2 DNA 二级结构约束	44
3.3.3 DNA 化学特性约束	45
3.3.4 DNA 子序列碱基组成约束	47
3.4 DNA 序列设计算法	47

3.5 基于热力学和汉明距离编码方法的分析比较	50
3.5.1 限制非特异性杂交的完备性比较	50
3.5.2 汉明距离的编码方法计算量分析	51
3.5.3 基于热力学的编码约束分析	53
3.5.4 核酸编码方法比较	56
第4章 DNA 二级结构预测和最小自由能模型	57
4.1 DNA 二级结构模型	57
4.2 动态规划 DNA 二级结构预测算法	60
4.3 改进的动态规划回溯树搜索	65
4.4 近邻热力学模型和最小自由能	71
4.4.1 Nearest-Neighbor 热力学模型	71
4.4.2 Watson-Crick 堆栈结构及其热力学模型	72
4.4.3 内单核苷酸错配结构及其热力学模型	74
4.4.4 内环结构及其热力学模型	75
4.4.5 凸环结构及其热力学模型	77
4.4.6 发卡结构及其热力学模型	78
4.4.7 悬挂末端结构及其热力学模型	79
4.4.8 多分支环和外环	80
4.4.9 DNA 二级结构预测例子	81
第5章 基于遗传算法的 DNA 平面伪结的预测算法	86
5.1 遗传算法概述	87
5.2 遗传算法的原理与方法	88
5.2.1 产生初始种群	88
5.2.2 根据目标问题构造适应度函数	88
5.2.3 选择适应度高的个体进行杂交和变异	88
5.2.4 若干次迭代后最高适应度值的为最优解	91
5.3 遗传算法和传统搜索算法的比较	92
5.3.1 梯度下降法	92
5.3.2 模拟退火算法	93
5.3.3 遗传算法	94
5.4 遗传算法预测含平面伪结的 DNA 二级结构	98
5.4.1 适应度函数	99

5.4.2 选择运算	101
5.4.3 交叉运算	101
5.4.4 变异运算	103
5.4.5 算法流程图	104
5.4.6 算法比较与分析	105
第 6 章 隐枚举核酸序列编码算法	108
6.1 隐枚举核酸序列设计算法	109
6.1.1 生成所有的 DNA 序列 n 元组	110
6.1.2 算法思路及步骤	111
6.2 DNA 编码约束转换隐枚举评估函数	113
6.2.1 转换汉明距离约束	115
6.2.2 转换相似度约束	116
6.2.3 转换 H-measure 约束及 3'端 H-measure 约束	117
6.2.4 转换反补汉明距离约束	119
6.2.5 转换自补汉明距离约束	120
6.2.6 转换单链二级结构约束	121
6.2.7 转换连续性约束	122
6.2.8 转换 GC 含量约束	124
6.2.9 转换解链温度约束	125
6.2.10 转换最小自由能约束	126
6.2.11 转换限定子序列约束和重叠子序列约束	127
6.2.12 转换模版序列约束	128
6.2.13 多约束的隐枚举评估函数	129
6.3 算法收敛性分析	130
6.4 同其他编码算法的比较	130
6.4.1 隐枚举算法和模板-映射算法的编码比较	130
6.4.2 隐枚举算法和模拟退火算法的编码比较	133
6.4.3 隐枚举算法和遗传算法的编码比较	134
6.4.4 隐枚举算法和最小子串算法的编码比较	135
6.4.5 隐枚举算法和多目标进化算法的编码比较	137
6.4.6 隐枚举算法和动态规划算法的编码比较	138

第 7 章 DNA 编码在图着色 DNA 计算中的应用	142
7.1 图顶点着色 DNA 计算算法	143
7.2 DNA 编码设计在图着色 DNA 计算中的应用	145
第 8 章 DNA 二级结构的平面嵌入算法	150
8.1 图	150
8.1.1 有向图	151
8.1.2 可平面图	151
8.1.3 顶点的度和图的度	152
8.1.4 连通图	153
8.2 ST-编号平面嵌入算法	154
8.2.1 ST-编号算法	154
8.2.2 平面嵌入	160
8.3 DNA 二级结构平面嵌入算法	163
8.3.1 布局算法	166
参考文献	170
后记	174

第 1 章 DNA 计算的产生与发展

DNA(deoxyribonucleic acid)计算是借助于分子生物技术进行计算的新方法,凭借极大的存储密度和高度并行性,DNA 计算为求解 NP 完全问题等复杂组合优化问题,提供了一种全新的途径,开创了广阔的应用前景。DNA 编码质量的优劣决定了 DNA 计算的效率,DNA 编码数量的多少决定了 DNA 计算可求解问题的规模,因此核酸编码是 DNA 计算研究中的重要课题。

随着 20 世纪 70 年代初期计算复杂性理论的形成,科学工作者发现并证明了大量来源于实际的组合优化问题是非常难求解的问题,即 NP 完全问题和 NP 难问题。20 世纪 80 年代初期,应运而生了一系列现代优化计算方法,如模拟退火算法、遗传算法、蚁群优化算法、粒子群算法和人工神经网络算法等。这类算法中的每一个算法都以自然、人类、生物的行为方式或物质运动形态为背景,经过数学抽象建立算法模型,通过电子计算机来求解组合优化问题。然而,这些基于仿生计算的优化计算方法不能在有效的时间中找到最优解,而用次优可行解来代替最优解;即使算法找到了最优解,算法本身也不能证明其是最优解。因此这一系列建立在模仿分子运动、遗传学、动物学、神经系统的现代优化算法并不能彻底解决复杂的组合优化问题。

DNA 计算是一种基于 DNA 分子的并行计算模型,是在计算科学和分子生物学的基础上发展起来的一个新颖而极具发展潜力的交叉学科,这种基于生物分子的计算模式在求解 NP 完全问题时显示出了极大潜力。

1994 年,美国南加州大学的 Adleman^[1]用 DNA 分子进行计算,求解出 7 个顶点的有向 Hamilton 路问题,成功地利用现代分子生物技术在 DNA 溶液的试管中进行了实验。这一研究成果很快引起了计算机、数学、分子生

物学等领域的科学家的极大兴趣。它的重要意义不仅在于算法和速度,更在于采用了一种全新的介质作为计算要件,以生物技术来实现电子计算机无法解决的复杂组合优化问题。随后,在 1995 年召开的第一届国际 DNA 计算会议上,来自各个国家和地区的 200 多名科学家共同探讨并充分肯定了 DNA 计算机的可行性。该领域的专家普遍认为:一旦 DNA 计算机研制成功,其运算量是目前的传统计算机望尘莫及的,DNA 计算是一个极具开发价值的研究领域。与传统的电子计算机相比,DNA 计算机具有三个突出的优点:高度并行性、海量的存储能力、极低的能耗。

然而,当前 DNA 分子的操作技术的发展远不如电子计算机中的集成电路成熟。在 DNA 计算过程中,DNA 分子间会发生非特异性杂交,导致实验出现错误甚至失败,极大降低 DNA 计算的可靠性和有效性。对 DNA 序列进行编码设计是 DNA 计算机研制中最为核心的问题。首先,编码质量的优劣决定了 DNA 计算的可靠性和有效性,直接影响着 DNA 计算能否按照预期的设计相互反应;其次,编码数量的多少决定了 DNA 计算求解问题规模的大小,与 DNA 计算机研究能否深入发展息息相关。所以,本书在详细分析编码质量和编码数量的主要影响因素的基础上,对 DNA 计算机中的编码问题及其算法进行了深入研究。

DNA 计算的基本思想:DNA 计算是以 DNA 相关的生物酶等作为基本材料,基于生化反应原理的一种新型的分子生物计算方法。利用 DNA 特殊的双螺旋结构和碱基互补配对规律进行信息编码,把要运算的对象映射成 DNA 分子链,然后在生物酶的作用下,按照一定的规则将原始问题的数据运算映射成高度并行的 DNA 分子链可控的生化反应。最后,利用分子生物技术如聚合酶链反应(PCR)、超声波降解、亲和层析、克隆、诱变、分子纯化、电泳、磁珠分离等,检测所需要的运算结果。DNA 计算的关键是将经过编码的 DNA 分子作为输入,在试管内或其他载体上经过一定时间完成可控的生物化学反应,最后提取出代表问题解的 DNA 分子。

1994 年 Adleman 求解 7 个顶点有向 Hamilton 路的实验解释了 DNA 计算过程的三个基本步骤,其有向图 G 如图 1.1 所示。

(1) 有向图的每一个顶点都编码为唯一的 DNA 分子,有向图的每一条边用相邻两顶点的 DNA 分子补链各取一半组成。

(2) 在一定的反应条件下,代表顶点和代表边的 DNA 分子相互杂交,

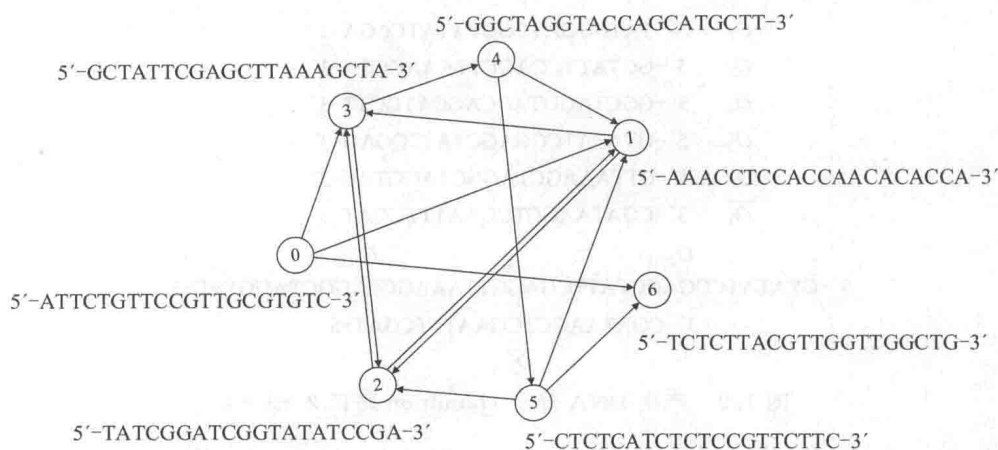


图 1.1 有向图 G 存在唯一 Hamilton 路 $v_0 \rightarrow v_1 \rightarrow v_2 \rightarrow v_3 \rightarrow v_4 \rightarrow v_5 \rightarrow v_6$

相互连接形成 DNA 分子,产生这个有向图的路径。

(3) 最后检测出代表该 Hamilton 路的 DNA 分子。具体计算过程如下。

(1) 解的产生。首先将图 1.1 中各顶点 v_i 用长度为 20 个碱基的单链 DNA (O_i) 来代替, v_i 和 O_i 一一对应。若顶点 v_i 和 v_j 之间存在边 $e(i, j)$, 再根据 O_i, O_j 来合成图中各条边 $O_{i \rightarrow j}$, $O_{i \rightarrow j}$ 同样由长度为 20 个碱基的单链 DNA 构成, 前 10 个碱基取 O_i 的 3' 端 10 个碱基, 后 10 个碱基取 O_j 的 5' 端 10 个碱基。这种 DNA 序列设计使得如果存在路径 $i \rightarrow j$ 和 $j \rightarrow k$, 通过添加 O_j 的补链 $\bar{O}_j$, 可以产生路径 $i \rightarrow j \rightarrow k$ 。

如图 1.2 所示, 以产生路径 $2 \rightarrow 3 \rightarrow 4$ 为例, O_2, O_3, O_4 分别为 20 个碱基组成的单链 DNA, 代表 v_2, v_3, v_4 三个顶点, $\bar{O}_3$ 是 O_3 的补链。单链 DNA $O_{2 \rightarrow 3}$ 和 $O_{3 \rightarrow 4}$ 代表该有向图的边 $e(2, 3)$ 和 $e(3, 4)$ 。其中 $O_{2 \rightarrow 3}$ 由 O_2 的 3' 端 10 个碱基和 O_3 的 5' 端 10 个碱基组成, $O_{3 \rightarrow 4}$ 由 O_3 的 3' 端 10 个碱基和 O_4 的 5' 端 10 个碱基组成。在 $O_{2 \rightarrow 3}$ 和 $O_{3 \rightarrow 4}$ 共同存在的 DNA 溶液中, 加入一定浓度的 $\bar{O}_3$, 则代表路径 $2 \rightarrow 3 \rightarrow 4$ 的双链 DNA 就产生了。依此类推, 当这一步连接反应包括了图 1.1 中所有的边 $e(i, j)$ 时, 图 1.1 中所有随机路径的 DNA 就可以在一歩之间全部产生了。

(2) 解的分离。在第(1)步中产生了各种路径, 首先需要选出起点和终点分别为 O_0 和 O_6 的路径。使用 O_0 和 $\bar{O}_6$ 作为引物, 利用 PCR 对第(1)步的结果进行扩增, 就能选择性地扩增那些从 O_0 至 O_6 的 DNA 路径。PCR 扩增

图 1.2 产生 DNA 分子 Hamilton 路径 $2 \rightarrow 3 \rightarrow 4$

完毕后, DNA 溶液中存在以 O_0 为起点、 O_6 为终点的不同长度的双链 DNA, 如图 1.3 所示。

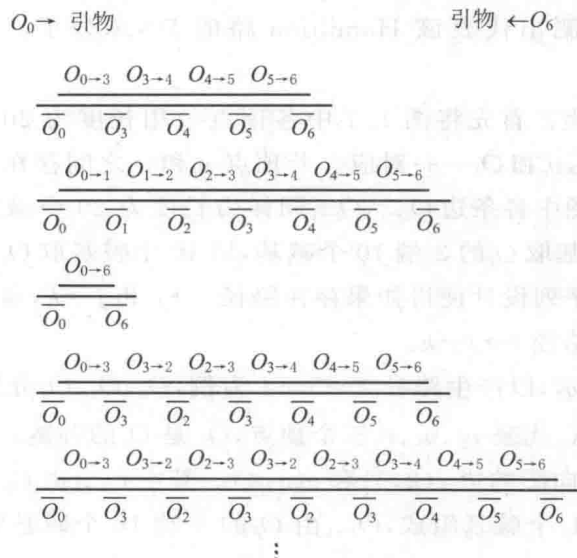


图 1.3 以 0 开始 6 结尾的所有路径

由于从 O_0 开始, 每增加一个顶点, 该双链 DNA 的长度就增加 20 个碱基。通过所有 7 个顶点的 DNA 路径, 对应的碱基个数必须为 140bp。通过凝胶电泳技术, 在电镜下观察得到 140 个碱基对应的谱带, 割下这条胶带, 浸入双蒸水中提取 DNA。

(3) 解的检测。第(2)步得到了长度为 140bp 的 DNA 溶液, 还要选出

其中经过每个顶点各一次的 DNA 路径,如图 1.4 所示。应用亲和纯化法,选出经过所有顶点至少一次的 DNA 路径。先将 DNA 双链解开双螺旋成单链 DNA,让其通过接有 O_1 补的磁珠,只有含有 O_1 序列的单链 DNA 才能由于互补链的结合而被滞留,这个操作分离出了经过 O_1 的 DNA。这样依次通过接有 O_2 补、 O_3 补、 O_4 补和 O_5 补的磁珠,最终得到了经过所有顶点至少一次的 Hamilton 路 DNA 分子。

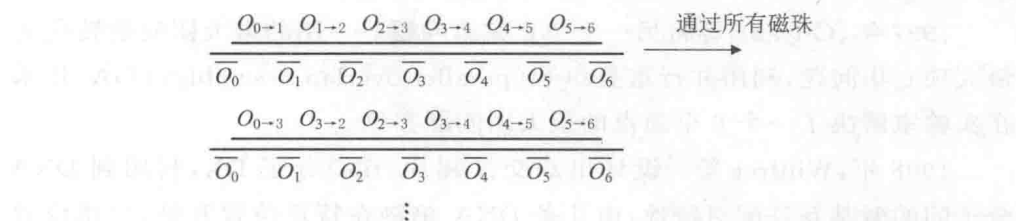


图 1.4 以 0 开始 6 结尾的等长 DNA 链通过磁珠

将第(3)步的产物再进行 PCR 扩增,通过 DNA 测序等方法,检测目标 DNA 是否存在即可。

1997 年 Garzon 等^[2]归纳了使用 DNA 分子解决计算问题的三个基本步骤:①问题编码(encoding),这个步骤将现实计算问题映射转换为 DNA 分子计算模型;②相互反应(interaction),这个步骤将设计好的 DNA 分子在一定温度、盐度和催化剂的作用下执行核心的生化计算;③解的提取(extraction),这个步骤利用生物检测技术将代表解的 DNA 分子提取出来,使纳米级的 DNA 分子在宏观世界可见。

1.1 国内外研究进展

Adleman 实验的成功标志着科学家试图利用生物大分子(DNA、RNA、蛋白质等)来解决电子计算机无法求解的 NP 完全问题和构建新型的计算机的设想取得了重要突破,使 DNA 计算成为一个备受关注的学科,也使构建新型的计算机——DNA 计算机成为可能。

随后,国内外学者在 Adleman 的分子计算模型基础上广泛开展了 DNA 计算的研究工作,建立了一系列求解 NP 完全问题的 DNA 计算模型,如

SAT、图着色、0-1 规划、背包问题等计算问题的 DNA 计算模型。

1995 年,受 Adleman 的启发,Lipton^[3]提出 DNA 计算可以解决另一种 NP 完全问题——可满足性问题(SAT)。Lipton 通过构造一个接触网络图,将 SAT 问题的解空间映射为通过接触网络图的起点和终点的所有哈密顿路。他首先利用 DNA 分子表示问题的所有可能解,然后利用生化反应删除非解。

1996 年,Roweis 等介绍了一种新的 DNA 计算模型——粘贴模型,并用此模型解决了最小集合覆盖问题和数据加密问题。

1997 年,Ouyang 等将另一个 NP 完全问题——图的最大团问题转化为最大独立集问题,利用并行重叠组装(parallel overlap assembly, POA)技术在实验室解决了一个 6 个顶点的最大团问题实例。

1998 年,Winfree 等^[4]设计出双交叉瓦片(tile)分子 DX,利用到 DNA 分子间的碱基互补配对特性,由几条 DNA 单链在特定位置互补,交错位置而形成分子结构,如图 1.5 所示。Winfree 发现以 DX 分子作为组件可以实现规则结构的自组装。

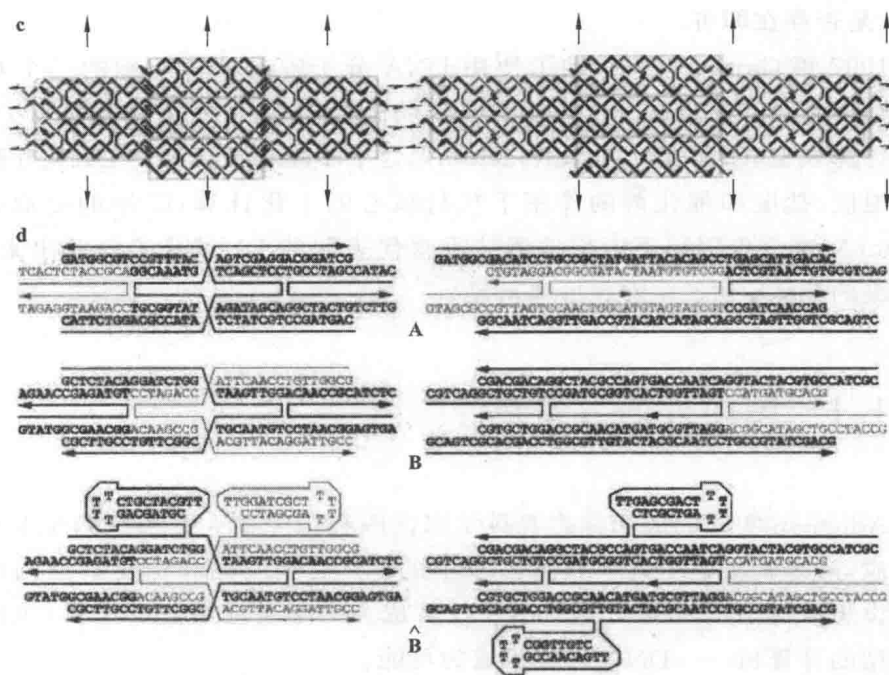


图 1.5 Winfree 自组装模型

值得关注的是,在 Winfree 的工作发表之前,DNA 分子普遍被理解为仅能两条单链绑定为双螺旋结构,而他构建的 DX 结构不仅超越了简单的一维线段,可利用多条 DNA 单链分子自组装成具有更大规模、更复杂的二维、三维 DNA 分子结构。该技术为后面一系列分子自组装高水平科研成果的研究奠定了基础。

1982 年,纽约大学的 Seeman^[5]就提出了利用 DNA 分子构造各种简单构件的思想。目前,已经能够在实验室制作出一些简单的构件,如三角形、长方形、圆形等,称为瓦片(tile)。Winfree 在其博士论文中首次对分子自组装进行了深入的研究,精心设计和选择了带有 3 个或 4 个黏性末端的三分支和交叉结构的瓦片 DNA 分子实现了抽象的拼图(tiling),并用于分子计算的自控制。结果表明:基于分子自组装行为的计算模型具有和图灵机等价的计算能力^[6]。通过对这些相同简单构件的自组装过程的研究,不仅可以进一步挖掘 DNA 计算解决优化问题的潜力,同时也可能实现对某些生物大分子间相互作用的调控。

2000 年,Labean 提出一种三交叉分子 tile 结构,较以往的 tile 类型在生化反应中更具刚性且更稳定。Mao 等在 *Nature* 发表基于三交叉 DNA 分子 tile 自组装完成四步累积异或运算。

2000 年,Wisconsin 大学的 Liu 等介绍了一种基于表面的 SAT 问题的解决方法。2001 年,Wu 对此方法进行了改进,使 DNA 计算机的发展向芯片化方向迈进了一大步。

2001 年,以色列的 Weizmann 研究所的 Benenson 等^[7]研制出由 DNA 分子和酶分子构成的生物计算机,该机器是一种可编程的、可自治地解决组合优化问题的有穷自动机。这种有穷自动机由 DNA 和操作 DNA 的生物酶构成。该自动操作的硬件由限制性核酸内切酶和连接酶构成,软件和输出由双螺旋输出编码,且编程相当于选择适当的软件分子。对这些成分的混合溶液,该自动操作处理由限制、杂交和连接循环的阶式蒸发器的输入分子,产生可检测的输出分子,这种输出分子对于自动操作的最终状态进行了编码,得到计算的结果。

2002 年,Braich 等^[8]给出了半自动化自组装 DNA 计算模型,并利用该模型解决了含有 20 个变元的 3 可满足性问题。他们利用所有解的组合构造 Sticker 模板,通过杂交对解进行检测。他们的思想与 Lipton 的思想基本