

基于自组装的 N 皇后问题 DNA 计算算法

吴 帆, 李肯立

(湖南大学信息科学与工程学院, 湖南长沙 410082)

摘 要: N 皇后问题是理论计算机科学中一个经典的 NP 难问题. 自 Adleman 首次运用 DNA 计算来解决 NP 问题以来, DNA 计算已成为计算机科学的研究热点之一, 现有 N 皇后问题的 DNA 计算机算法多基于粘贴和剪接模型, 存在生化操作复杂度和实验误差较高等问题. 本文提出了一种基于 DNA 自组装模型来求解 N 皇后问题的 DNA 计算方法. 算法通过减少实验操作步骤数, 降低了生化解的错误率. 算法使用的 tiles 分子块种类为 $O(n^2)$, 生化操作复杂性为 $O(1)$, 其中 n 为皇后的个数. 与求解 N 皇后问题的其它 DNA 算法的对比分析表明, 本算法可提高生化解的准确性, 降低算法生化实验的复杂度, 具有良好的易操作性.

关键词: DNA 计算; 自组装模型; N 皇后问题; tile 模型

中图分类号: TP3 **文献标识码:** A **文章编号:** 0372-2112 (2013) 11-2174-07

电子学报 URL: <http://www.ejournal.org.cn> **DOI:** 10.3969/j.issn.0372-2112.2013.11.010

An Algorithm in Tile Assembly Model for N Queen Problem

WU Fan, LI Ken-li

(College of Electrical and Information Engineering, Hunan University, Changsha, Hunan 410082, China)

Abstract: DNA computing employs molecule manipulation to solve NP complete problems that can not be solved using traditional Turing machine. With the deep studying of DNA computing, we found that DNA computation suffers from relatively high error rates. How to decrease error rates has become an important part of DNA computing. This paper present DNA self-assembly model which decreases error rates through decreasing experiment operations. A new N Queen problem algorithm based on the self assembly model is proposed. The proposed algorithm needs $O(n^2)$ types of tiles and the complexity of experiment operations is $O(1)$. Obviously, this algorithm significantly reduces the complexity of the experiment, thus improving the accuracy of experimental results.

Key words: DNA-based computing; DNA self-assembly model; N queen problem; tile model

1 引言

自组装 (self-assembly) 是系统的构成元素在不受外力的介入下, 自行聚集、组织成规则结构的现象. 自组装过程通常会将系统从一个无序的状态转化成一个有序的状态. 1963 年, Wang 从数学角度以 tiles 分子块作为计算指令, 指出 tiles 可以通过自组装过程在二维平面形成周期性格局, 且证明了该过程具有图灵等价性^[1].

1994 年, Adleman 运用 DNA 分子解决 7 节点的 Hamilton 路径问题^[2], 首次提供了一种以 DNA 计算来解决 NP 难问题, 并且证明了运用 DNA 分子进行计算的可行性. 随后, 许多学者研究利用 DNA 计算来求解 NP 问题^[3,4], 文献[5,6]提出利用粘贴模型来求解最大团和三维匹配问题. 1998 年, Winfree 和 Seeman 基于 Wang Tiling

的思想, 提出通过 DNA Tile 分子自组装的过程来实现计算的思想, 并证明 DNA Tile 自组装计算具有图灵等价性^[8,9]. 2008 年, Mao 应用 DNA 单链进一步组装形成了不同多面体结构^[10], 文献[11]提出基于三维结构的 tile 分子模型. 由于基于自组装模型的 DNA 计算比其它模型存在操作复杂度较低, 出错率比较低的优点, 近年来, 不少学者提出了基于自组装模型的求解 NP 问题 DNA 计算算法^[12].

N 皇后问题是理论计算机科学中一个经典的 NP 难问题. 文献[15]提出了基于粘贴模型 N 皇后问题的 DNA 计算机算法, 尽管该算法在多项式时间内求得问题的解, 然而利用粘贴、剪接模型等来求解 NP 问题时, 虽然这些算法能够在多项式时间内求得 NP 的解, 但是其生化操作具有较高的复杂度. 生化实验操作步骤较多, 相

应带来的试验误差也较大.本文基于 DNA 自组装模型,提出了 N 皇后问题的一种新的 DNA 计算算法.该算法不仅可以在线性时间内精确求解 N 皇后问题的解,且生化实验复杂度低,准确性高,具有良好的实验操作性.

2 DNA 自组装计算模型

2.1 自组装模型的 DNA Tile 二维平面结构

DNA tile 分子是一簇 DNA 交叉分子,每一个分子都有一个黏贴末端,可以与具有互补黏贴末端的 Tile 分子嵌套结合,逐步形成 DNA 分子网格结构. DNA 链根据碱基互补配对规则自组装可形成一维、二维、三维结构. Adleman 首次构建的 DNA 计算是利用 DNA 链组装形成的一维结构.但是其生化实验步骤会随着问题规模的增加成线性增加^[12],这就给生化实验带来了挑战.为了克服这种不足,Winfree 等人提出二维的自组装计算模型且设计了自组装单元 DNA 双交叉分子 DX^[9]. DX 指在螺旋域之间包含两个交叉链接的 DNA 分子结构^[16],实验发现可成功构建五种 DX 结构,但发现有 DAO(double-crossover antiparallel odd)和 DAE(double-crossover antiparallel even)这两种结构是稳定的^[9].本文采用二维结构的 DAE 交叉分子来作为 DNA tile 分子(如图 1),同时将其抽象为具有四个方向的 tile 矩形块(如图 2).

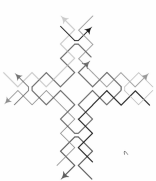


图1 DAE交叉分子

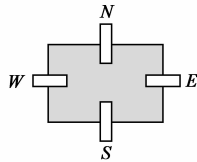


图2 抽象tile分子块

2.2 自组装计算模型的数学描述

根据 Wang 的 Tile 理论模型的形式化描述以及 Yuriy Brun 引入的 DNA 的 tile 自组装模型^[17],本文将基于 DAE 结构的自组装模型表示为 $S = \langle T, g, \tau \rangle$ 其中 T 为 DAE Tile 分子块的集合, g 为能量函数, τ 为常数. 相应的符号描述如下:

(1) Σ : DAE 块中所有粘性末端所表示的符号的集合,即所要参与计算的符号的集合.

(2) $D' = \{N, E, S, W\}$: 分别表示 DAE 块的北端,东端,南端,西端四个位置与方向.

(3) $\{\sigma_N, \sigma_E, \sigma_S, \sigma_W\} \in \Sigma^4$: DAE 块的形式化描述,表示 DAE 块的四个位置分别所表示的符号.

(4) $(x, y) \in Z^2$: 表示 DAE 块在二维平面上的位置.

(5) $D = \{DN, DE, DS, DW\}$: 表示 4 个方向函数,且满足有如下特点:

(5.a) $\forall (x, y) \in Z^2 DN(x, y) = (x, y + 1), DE(x, y) = (x + 1, y), DS(x, y) = (x, y - 1) DW(x, y) = (x - 1, y).$

(5.b) 若位置 (x, y) 与 (x', y') 相邻,则 $\exists d \in D$, 使得 $d(x, y) = (x', y')$.

(6) $bd_d(t) \in \Sigma$, 其中 $t \in T, d \in D'$: 这是一个函数,用来得到 DAE 分子块 t 在位置 d 方向的粘性末端所表示的符号.

(7) 函数 $g: \Sigma \times \Sigma \rightarrow R$, 此函数为能量函数,定义如下.

(7.a) $\forall \sigma \in \Sigma$, 则 $g(\text{null}, \sigma) = 0$

(7.b) 若 $g(\sigma, \sigma') = 0$ (其中 $\sigma, \sigma' \in \Sigma$), 则 $\sigma \neq \sigma'$

(7.c) $\forall \sigma \neq \text{null}$, 有 $g(\sigma, \sigma) = 1$

(7.d) $\forall \sigma \neq \sigma'$, 有 $g(\sigma, \sigma') = 0$

(8) 对于格局 C , 存在函数 $A_C(x, y): Z \times Z \rightarrow T$, 此函数有如下特点:

(8.a) 若 $(x, y) \in Z^2$, 且 $A_C(x, y) \neq \text{empty}$, 则 $(x, y) \in C$

(8.b) 若只有有限个不同的 $(x, y) \in C$, 则 C 为一个有限集

(9) 若一个 DAE 块 t 能在格局 C 的 (x, y) 位置结合产生一个新的格局 C' , 需要满足以下条件:

(9.a) $(x, y) \notin C$

(9.b) 若 $d' \in D, d$ 与 d^{-1} 为相对的方向, $\sum_{d' \in D'} g(bd_{d'}(t), bd_{d^{-1}}(A_C(d'(x, y)))) \geq \tau$

(9.c) $\forall (u, v) \in Z^2, (u, v) \neq (x, y)$, 则 $A_C(u, v) = A_C(u, v)$

(9.d) $A_C(x, y) = t$

3 自组装 N 皇后问题的 DNA 算法

3.1 N 皇后问题的 DNA 计算算法框架

N 皇后问题就是在 $N \times N$ 的棋盘上,放置 N 个皇后,要求每一行每一列以及每一对角线上均只能有一个皇后.基于自组装求解 N 皇后问题的 DNA 算法框架为:

第一步,构建表示 N 皇后问题解空间的初始子系统.

第二步,通过 DNA 自组装过程生成求解子系统,求出 N 皇后问题的解.如图 3 所示,第一个部分用来计算棋盘中每一列是否存在冲突,第二部分用来判断其顺时针对角线上是否只有一个皇后,第三部分用来判断其逆时针对角线上皇后是否存在冲突;

第三步,利用检测子系统,从求解子系统中找出问题的解.通过提取带有荧光标记的检测分子(如图 6),得到了 N 皇后问题的解.

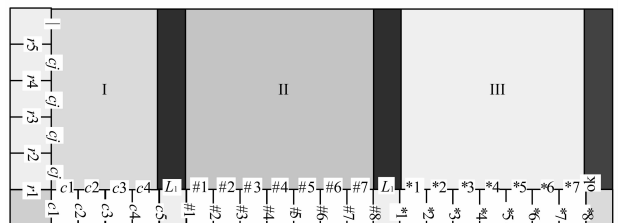


图3 最终格局的DNA分子团示意图

3.2 DNA tile 分子块的设计

3.2.1 初始分子的设计

利用自组装模型求解 N 皇后问题时,首先需对该问题编码.本文设计的初始分子 Γ_{R_i} (如图 4(h)) ($1 \leq i \leq N$) 分子用来实现此目标. Γ_{R_i} 分子的 $\sigma_S = r_i, \sigma_N = r_i + 1, \sigma_E = c_j, \sigma_W = \text{null}$, 其中 $1 \leq i \leq N, 1 \leq j \leq N$. 如果 i 表示皇后所在棋盘的行号, j 表示皇后所在棋盘的列号, 则一个 Γ_{R_i} 分子就表示皇后在棋盘上的位置, 将 N 个 Γ_{R_i} ($1 \leq i \leq N$) 分子根据 i 的大小从小到大排成一列, 就表示了 N 皇后在棋盘的一个分布, 且该分布表示棋盘上的每一行均只有一个皇后.

为求出 N 皇后问题的解, 还需计算每一个分布是否满足约束条件. 因此设计初始分子 Γ_{c_i} (如图 4(b)) ($1 \leq i \leq N$) 来编码棋盘的列号, 其 $\sigma_N = c_i, \sigma_E = c_{i+1}, \sigma_W = c_i, \sigma_S = \text{null}$, 其中 $1 \leq i \leq N, i$ 为棋盘的列号; 初始分子 $\Gamma_{\#i}$ (如图 4(c)) 用来编码棋盘的顺时针对角线号, 其 $\sigma_N = \#i, \sigma_E = \#i + 1, \sigma_W = \#i, \sigma_S = \text{null}$; 初始分子 Γ_{*i} (如图 4(d)) 用来编码棋盘的逆时针对角线号, 其 $\sigma_N = *i, \sigma_E = *i + 1, \sigma_W = *i, \sigma_S = \text{null}$. 在形成初始子系统的过程中, 这些分子按照一定的顺序粘贴在一起形成一行, 构建了三个不同的区域 (如图 3), 分别计算在棋盘的列上、顺时针对角线以及逆时针对角线是否存在冲突. 另外, 初始连接分子 Γ_1 和 Γ_2 (如图 4(e), (f)) 用来实现三个区域之间的连接, 其中 Γ_1 的 $\sigma_N = L_1, \sigma_E = c_{N+1}, \sigma_W = \#1, \sigma_S = \text{null}$, Γ_2 的 $\sigma_N = L_1, \sigma_E = \#2 \times N, \sigma_W = *1, \sigma_S = \text{null}$. 初始边界分子 Γ_S, Γ_3 和 Γ_{R_i} (如图 4(a), (g), (i)), 主要用来结束在某方向上的自组装行为, 其中 Γ_S 的 $\sigma_N = r_1, \sigma_E = c_1, \sigma_S = \text{null}, \sigma_W = \text{null}$; Γ_3 的 $\sigma_N = OK, \sigma_W = *2 \times N, \sigma_E = \text{null}, \sigma_S = \text{null}$ 以及 Γ_{R_i} 的 $\sigma_S = r_n, \sigma_E = 1, \sigma_W = \text{null}, \sigma_N = \text{null}$. 对于 $n \in N$ 皇后问题, 初始分子集中 tile 分子类型的个数为 $n^2 + 5n + 3$.

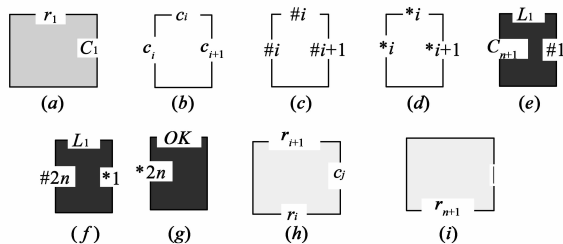


图4 初始分子集

3.2.2 规则分子的设计

四类规则分子用来计算初始子系统中的一个初始格局是否满足相应的约束条件: 第一类为列约束计算规则分子, 见图 5 的 (a)(b)(c). 其中 (a) 为紫色分子, $\sigma_W = \sigma_E = \sigma_S = c_i, \sigma_N = i, 1 \leq i \leq N$ 表示棋盘第 i 列上正好找到一个皇后. (b) 为棕色分子, $\sigma_W = \sigma_E = c_i, \sigma_N =$

$\sigma_S = c_j, i \neq j$ 其中 $1 \leq i, j \leq N$, 表示棋盘第 j 列还未找到皇后. (c) 为淡绿色分子, $\sigma_W = \sigma_E = c_i, \sigma_N = \sigma_S = j, i \neq j$, 其中 $1 \leq i, j \leq N$, 表示棋盘上第 j 列已经找到一个皇后;

第二类顺时针对角线约束计算规则分子, 见图 5 中 (e)(f)(g). (e) 为深红色分子, $\sigma_W = \sigma_E = \sigma_S = c_i, 1 \leq i \leq 2 \times N - 1$, 表示棋盘第 i 号顺时针对角线上正好找到一个皇后. (f) 为绿色分子, 其中 $\sigma_W = \sigma_E = \#i, \sigma_N = \sigma_S = \#j, i \neq j, 1 \leq i, j \leq 2 \times N - 1$, 表示棋盘第 j 号顺时针对角线上还未找到一个皇后. (g) 为淡蓝色分子, $\sigma_W = \sigma_E = \#i, \sigma_N = \sigma_S = j, i \neq j, 1 \leq i, j \leq 2 \times N - 1$, 表示表示棋盘上第 j 号顺时针对角线已经找到了一个皇后;

第三类逆时针对角线约束计算规则分子, 如图 5 (i)(j)(k). (i) 为紫色分子, $\sigma_W = \sigma_E = \sigma_S = *i, 1 \leq i \leq 2 \times N - 1$, 表示棋盘第 i 号逆时针对角线上找到一个皇后. (j) 为深绿色分子, 其中 $\sigma_W = \sigma_E = *i, \sigma_N = \sigma_S = *j, i \neq j, 1 \leq i, j \leq 2 \times N - 1$, 表示棋盘第 j 号顺时针对角线上还未找到皇后. (k) 为黄色分子, $\sigma_W = \sigma_E = *i, \sigma_N = \sigma_S = j, i \neq j, 1 \leq i, j \leq 2 \times N - 1$, 表示棋盘上第 j 号顺时针对角已经找到皇后.

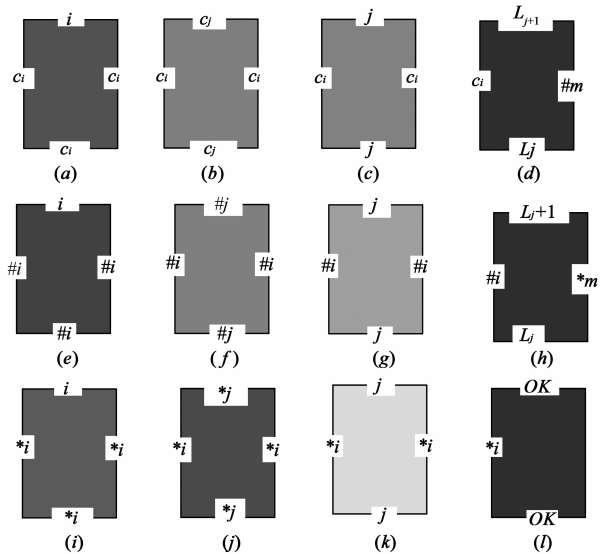


图5 规则分子集

第四类是连接分子, 用来连接列判断区域和对角线判断区域. 它们均为蓝色分子, 如图 5 (d)(h)(l), 图 5 (d) tile 分子的 $\sigma_W = c_i, \sigma_E = \#m, \sigma_N = L_j, \sigma_S = L_{j+1}$, 其中 i 表示行号, j 表示列号, m 表示棋盘 (i, j) 的顺时针对角线号, 且 $1 \leq i, j \leq N, m = i + j - 1$; 图 5 (h) 分子块的 $\sigma_W = \#i, \sigma_E = *m, \sigma_N = L_j, \sigma_S = L_j, \sigma_S = L_{j+1}$, 其中 i, j 所表示的含义同, m 表示棋盘 (i, j) 的逆时针对角线号, 且 $1 \leq i, j \leq N, m = i - 2 \times j + n + 1$; 图 5 (l) 分子块的 $\sigma_E = *2 \times N, \sigma_W = \text{null}, \sigma_N = OK, \sigma_S = OK$.

3.2.3 检测分子的设计

检测子系统需要从求解子系统提取问题的解,因此设计了如下几类检测分子:第 1 类是检测 N 皇后在棋盘上的每一列是否存在冲突,这类分子如图 6(a)~(e),其中(a)为列检测分子 $TCe_i, 1 \leq i \leq N$;第 2 类是用来检测 N 皇后在顺时针对角线是否存在冲突,如图(c),为顺时针对角线检测分子 $T\#e_i, 1 \leq i \leq 2 \times N - 1$;第 3 类用来检测 N 皇后在逆时针对角线是否存在冲突,如(d),为逆时针对角线检测分子 $T * e_i, 1 \leq i \leq 2 \times N - 1$;第 4 类为蓝色分子,用来连接各部分的检测分子,如(b);第 5 类为红色的 T_{OK} 分子,该分子将标记银光粉,它用来提取 N 皇后问题的解。

3.3 算法步骤

利用本文设计的 DX 分子,结合基因工程相应的生化操作,提出了基于自组装模型的 N 皇后问题 DNA 计算算法。

算法 1 基于自组装模型的 N 皇后问题 DNA 计算算法(N-queen problem algorithm)

输入:tile 分子集 T_1, T_2, T_3 , 他们分别为初始分子集,规则分子集,检测分子集,

输出: N 皇后问题的解

Procedure:N-queen-algorithm(G, T_1, T_2, T_3)

(1) $T = \text{Anneal}(T_1)$, 将初始分子经退火,生成初始系统;

(2) $T = ((T, T_2, T_3)$, 将规则和检测分子放入初始系统中;

(3) $T = \text{Anneal}(T)$, 经退火反应,生成求解和检测系统;

(4) 使用 AFM 观察具有荧光标记的最终格局分子块;该分子快所代表的布局就是 N 皇后问题的一个解。

(5) $T_{\text{end}}' = \text{Extract}(T)$, 提取在(6)中所观测到的分子块,将其放入试管 T_{end}' , 通过解析分子团所包含的 N 皇后的布局,就得到 N 皇后问题的解;

4 算法分析

引理 1 $\sum = \{r_i, c_i, *j, \#j, i, L_1, L_2, OK, \mid\}$,

其中 $0 < i \leq N, 0 < j \leq 2 \times N + 1, \Gamma$ 为 3.3.1 中设计的初始分子集,令 $\tau = 1, g = 1$. 则对 Γ 中的初始分子进行 PCR 聚合酶链式反应,充分退火后生成初始子系统 S , 其中 $S = \langle \Gamma, g, \tau \rangle$, 则初始子系统 S 中包含的初始格局形状为 L 型,如图 7, 每一个初始格局均可表示 N 皇后在棋盘上的一个分布,且初始系统所包含的全部初始格局能包含 N 皇后问题的所有解。

证明: 设初始分子 Γ_S 在初始格局中的位置 $(0, 0)$, 因 Γ_S 的 $\sigma_w = \sigma_s = \text{null}$, 因此初始子系统的自组装过程只能沿着 Γ_S 分子的 N 端和 E 端方向展开. 考察 Γ_S 分子 N 端的自组装过程: 因其 $\sigma_N = r_1$, 而 $\tau = 1, g = 1$. 据自组装计算模型的第 9 条(见本文 2.2)以及初始分子

集可知,初始格局 $(0, i)$ 位置上只能为 Γ_{Ri} 分子. 因此初始格局第 0 列的 $(0, 1)$ 到 $(0, n)$ 位置上的分子相应为 $\Gamma_{R1} \sim \Gamma_{Rn}$. 又 Γ_{Rn} 的 $\sigma_N = r_{n+1}$. 所以 $(0, n+1)$ 位置上为边界分子 Γ_{Re} . 由于 Γ_{Ri} 分子可表示棋盘的 i 行有一个皇后,因此初始格局的第 0 列便可表示 N 个皇后在 $N \times N$ 棋盘上的一个分布,且每一行有且只有一个皇后。

再看 Γ_S 分子 E 端: 由于始端分子 Γ_S 的 $\sigma_w = c_1$, 故只有 Γ_{Ci} 才能粘贴在 $(i, 0) (0 < i \leq N)$ 上. $(N, 0)$ 位置上为 Γ_{Cn} 分子, 其 $\sigma_w = c_{N+1}$, 因此初始格局中 $(N+1, 0)$ 位置上只能为连接分子 Γ_1 . 又 Γ_1 的 $\sigma_w = \#_1$, 所以 Γ_1 右边的分子只能为 $\Gamma_{\#1}$. 因此, $\Gamma_{\#i} (0 < i \leq 2 \times N - 1)$ 分子将从左至右依次按照 i 的大小粘贴在初始格局中的 $(N+2, 0)$ 到 $(3 \times N, 0)$ 位置上. 同理通过连接分子 $\Gamma_2, \Gamma_{*i} (0 < i \leq 2 \times N - 1)$ 分子也将从左至右依次按照 i 的大小粘贴在初始格局中的位置 $(3 \times N + 2, 0)$ 到 $(5 \times N, 0)$ 上. 且边界分子 Γ_3 在位置 $(5 \times N + 1, 0)$ 上。

经上述分析可知,初始子系统中生成的初始格局为 L 型,第 0 列由 Γ_{Ri} 分子组成,只要初始系统中有足够多的初始格局,其就可包含 N 皇后问题的所有解. 第 0 行包括三个部分,第一部分由 Γ_{Ci} 分子组成,表示棋盘的第 1 到第 N 列;第二部分由 $\Gamma_{\#i}$ 分子组成,表示棋盘第 1 到第 $2 \times n - 1$ 的顺时针对角线;第三部分由 Γ_{*i} 型分子组成,表示棋盘第 1 到第 $2 \times n - 1$ 的逆时针对角线. 每部分之间用连接分子连接. 初始子系统 S 可形式化描述为如下: ① $A_S(0, 0) = \Gamma_S$; ② 对于 $0 < i \leq N$, 有 $A_S(0, i) = \Gamma_{Ri}$; ③ $A_S(0, N+1) = \Gamma_{Re}$; ④ 对于 $0 < j \leq N$, 有 $A_S(j, 0) = \Gamma_{Rj}$; ⑤ $A_S(N+1, 0) = \Gamma_1$; ⑥ 对于 $N+1 < j \leq 3 \times N, k = j - N - 1$, 有 $A_S(j, 0) = \Gamma_{\#k}$; ⑦ $A_S(3 \times N + 1, 0) = \Gamma_2$; ⑧ 对于 $3 \times N + 1 < j \leq 5 \times N, k = j - 3 \times N - 1$, 有 $A_S(j, 0) = \Gamma_{*k}$; ⑨ $A_S(5 \times N + 1, 0) = \Gamma_3$. (证毕)

引理 2(唯一性) 设自组装系统 $S_n^2 = \langle T, g_n^2, \tau_n^2 \rangle$, 其 $g_n^2 = 1, \tau_n^2 = 2$, 如果 $\forall t \in T$, 只要 t 的 (σ_S, σ_N) 对唯一, 那么由引理 1 生成的每一个初始格局与规则分子通过自组装反应后, 只能生成唯一的最终格局。

证明: 考察规则分子集 T 可以发现, 每一个分子块它的 S 端与 N 端的粘性末端符号对都是唯一, 因此据文献[21]中的定理 2.2 可知, 每一个初始格局都能唯一确定一个最终的格局. (证毕)

定理 1 对任意 $N \in N$, 如果 $\sum = \{r_i, c_i, *j, \#j, i, L_1, L_2, OK, \mid\}$, 其中 $0 < i \leq N, 0 < j \leq 2 \times N - 1$, 设初始分子集、规则分子集(如图 5)和检测分子集(如图 6)为 T , 设 $g_2 = 1, \tau_2 = 2, g_2$ 表示能量函数, τ_2 为热力学温度常数, 则系统 $S_2 = \langle T, g_2, \tau_2 \rangle$ 经求解子系统和检测子系统后能得到 N 皇后问题的解。

证明: 因 $g_2 = 1, \tau_2 = 2$, 故在自组装过程中, 初始格

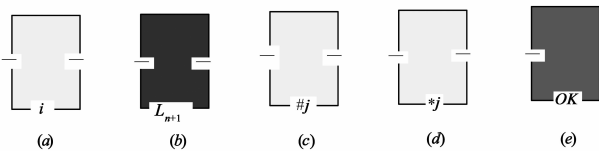


图6 检测分子集

局中能粘贴规则 *tile* 分子块的第一个位置只能是 $(1, 1)^{[21]}$. 从位置 $(1, 1)$ 开始, 由下至上, 从左至右依次粘贴分子形成新的格局, 最终在格局的边界粘贴检测分子终止自组装过程.

初始格局包含三个区域(如图 3), 先考察初始格局中第一个区域, 如果其第 i 列 $(1 \leq i \leq N)$ 的 $(i, N+1)$ 位置上粘贴有检测分子, 则在该列中有且必有一个紫色 *tile* 分子, 也就表示棋盘的第 i 列有且只有一个皇后, 否则就表示在棋盘的第 i 列有多个或 0 个皇后. 因此, 只要最终格局在 $(N, N+1)$ 位置上粘贴有检测分子 TCe_N , 就表示该初始格局所编码的 N 皇后分布在棋盘每一行, 每一列都不存在冲突.

再考察初始格局的第二和第三区域: 这两个区域用来计算顺时针和逆时针对角线上的皇后是否存在冲突. 因为初始格局中只编码皇后所在位置的行号和列号, 然而第二部分和第三部分要判断棋盘的对角线是否存在冲突, 因此利用本文设计的连接分子可将初始格局所代表的皇后所在的行、列号转换为其对角线号:

考察初始格局的第 $N+1$ 列, 它连接第一部分和第二部分(见图 3). 从规则分子集得知, 除了连接分子之外, 其余所有的分子其 $\sigma_E = \sigma_W$, 因此有:

$$bd_W(A_C(N, i)) = bd_W(A_C(0, i)) = C_j, 1 \leq i \leq N, 1 \leq j \leq N.$$

所以在位置 $(N+1, i)$ 上的分子有: $bd_E(A_C(N+1, i)) = C_j, bd_S(A_C(N+1, i)) = L_i, bd_W(A_C(N+1, i)) = \#_m, 1 \leq i \leq N, 1 \leq j \leq N$, 据连接分子(见 3.2.2)可知: $m = i + j - 1$. 因此如果 i 表示棋盘的第 i 行, 而 j 表示棋盘的第 j 列, 那么 m 表示棋盘 (i, j) 所在的顺时针斜对角线号. 由上推理, 如果第二部分的每一列均粘贴了检测分子, 那么 N 皇后在顺时针斜对角线上就不存在冲突; 否则如果某一列上没有粘贴检测分子, 如图 8(a), 则说明在该顺时针斜对角线上必定存在多个皇后. 同理, 对于第三部分, 通过连接分子可以将初始格局所代表的 N 个皇后所在的行列号转换为逆时针对角线号, 然后利用相应的规则分子和检测分子可以计算各皇后是否逆时针对角线上存在冲突. 如果第三部分的每一列均粘贴有检测分子, 表明各皇后在逆时针对角线上不存在冲突. 否则, 皇后在逆时针对角线上存在冲突.

通过以上分析可知: 如果在最终格局 S 上有 $A_S(5 \times N + 1, N + 1) = T_{ok}$, T_{ok} 分子为最终检测分子, 则说明该格局所代表的 N 皇后布局在每一列, 每一个对角线

上都最多有 1 个皇后, 即其为 N 皇后问题的一个解. 因为 T_{ok} 上已经标记有荧光素, 因此只需要使用 AFM 找出所有具有荧光标记的分子块, 便可以求出 N 皇后问题的解.(证毕)

定理 2 系统 S_n^2 的自组装时间复杂度是为 $O(n)$

证明 由定理一可知, 每一初始格局只能生成一个最终格局. 对于 n 个皇后, 最终格局最多能粘贴 $(n+1) \times (5 \times n + 1)$ 个 *tile* 分子, 因此根据文献[21]中的定理 2.3 可知, 生成最终格局的自组装时间为 n 的线性时间, 其时间复杂度为 $O(n)$. (证毕)

定理 3 系统 S_n^2 自组装过程中需要的分子块的类型数量为 $O(n^2)$.

证明 对于 N 皇后问题, 设皇后的个数为 n . 则本文设计的初始分子的类型个数为 $n^2 + 5n + 3$, 规则分子的类型个数应为 $6 \times n^2 + 11 \times n + 1$, 检测分子的类型个数为 $5 \times n + 5$, 自组装过程中需要的分子块的类型数量为 $O(n^2)$. (证毕)

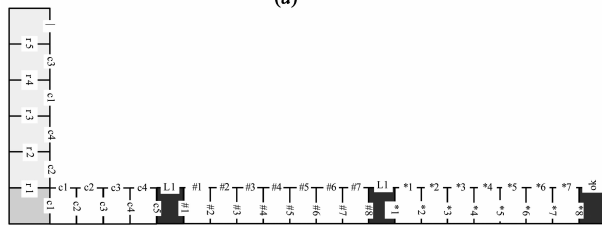
5 仿真实验分析

为了测试本文中算法的有效性, 我们利用自组装仿真软件 *xgrow*^[7, 19, 23], 模拟实现了 4 皇后问题本文的 *dna* 自组装算法.

利用本文 3.2.1 节中的方法设计了 4 皇后问题的初始分子, 根据算法第 1 和 2 步, 通过 *xgrow* 生成包含 4 皇后问题解空间的初始系统, 图 7(a) 为 *xgrow* 软件模拟自组装生成的一个初始格局, 图 7(b) 为图 7(a) 的符号表示. 该初始格局的最左边列表示 4 个皇后在棋盘上的位置分布, 该分布用一个 4 元组 $(2, 4, 1, 3)$ 表示, 即表示 4 个皇后分布在棋盘位置 $(1, 2), (2, 4), (3, 1)$ 和 $(4, 3)$ 上.



(a)



(b)

图7 4皇后问题的一个初始格局

再据算法第 3 步到第 5 步, 将规则分子以及检测分子与初始子系统经过充分退火反应后将得出最终格

局.如图 8 为 xgrow 模拟生成的两个典型最终格局.分析图 8 可知:(a),(b)代表 4 个皇后在棋盘上的分布分别为(1,3,2,4),(2,4,1,3).对于图 8(a),因该布局在棋盘第 4 号顺时针对角线上有两个皇后,因此自组装过程在最终格局的第 2 部分的第 4 列终止;对于图 8(b),因为 4 个皇后在每一行,每一列和两个对角线上均不存在两个皇后,因此红色的检测分子粘贴在最终格局上,表示这个最终格局均为 4 皇后问题的一个解,其中图 9 为图 8(b)的抽象表示.

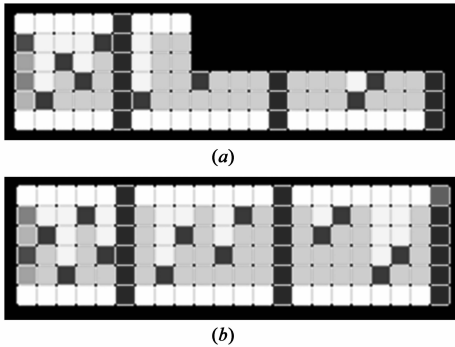


图 8 xgrow 生成的 4 皇后问题最终格局

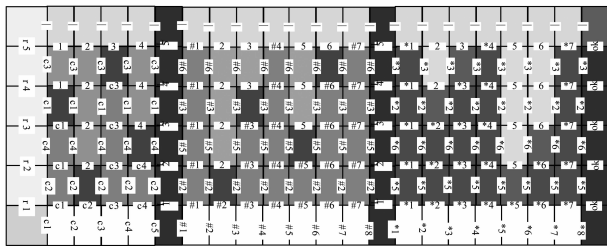


图 9 图 8(b)的最终格局表示图

最后根据算法 1 的第 6 步,将提取所有带有红色荧光的 tile 分子的最终格局,这些格局所代表的就是 N 皇后问题的解.由此求得四皇后问题的解为(2,4,1,3)和(3,1,4,2).

6 结论

自从 Adleman 首次运用 DNA 计算来解决 NP 问题以来,有不少学者提出 N 皇后问题的 DNA 计算机算法.许多学者是利用粘贴,剪接模型等,虽然这些算法能够在多项式时间内求得 N 皇后问题的解,但是其生化操作具有较高的复杂度.生化实验操作步骤越多,相应带来的试验误差也就越大.无法保证生化解的准确性.本文利用 DNA 自组装模型,根据 DNA 链形成的 tiles 的自组装特性以及自组装计算模型,提出了一种 N 皇后问题的 DNA 自组装算法.算法的生化操作复杂度为常数级,且提高了生化解的准确性,具有良好的易操作性.

参考文献

- [1] Wang H. Dominoes and the AEA case of the decision problem [A]. Proceedings of the Symposium in the Mathematical Theory of Automata. Brooklyn [C]. New York: Polytechnic Press, 1963. 23 – 55
- [2] Adleman L M. Molecular computation of solutions to combinatorial problems [J]. Science, 1994, 266(11): 1021 – 1023.
- [3] 刘文斌,朱翔欧,王向红,等. DNA 计算的研究进展 [J]. 电子学报, 2006, 34(11): 2053 – 2057.
Liu Wen-bin, Zhu Xiang-ou, Wang Xiang-hong et al. Progress in the research of DNA computing [J]. Acta Electronica Sinica, 2006, 34(11): 2053 – 2057. (in Chinese)
- [4] 许进,谭刚军,范月科,等. DNA 计算机原理、进展及难点 (IV): 论 DNA 计算机模型 [J]. 计算机学报, 2007, 30(6): 881 – 893.
Xu Jin, Tan Gang-Jun, Fan Yue-Ke, et al. DNA computer principle, advances and difficulties (IV): On the models of DNA computer [J]. Chinese Journal of Computers, 2007, 30(6): 881 – 893. (in Chinese)
- [5] 范月科,强小利,许进. 图的最大团与最大独立集粘贴 DNA 计算模型 [J]. 计算机学报, 2010, 33(2): 305 – 310.
Fan Yue-Ke, Qiang Xiao-li, Xu Jin. Sticker model for maximum clique problem and maximum independent set [J]. Chinese Journal of Computers, 2010, 33(2): 305 – 310. (in Chinese)
- [6] 周旭,李肯立,乐光学,杨志邦. 一种基于分治的三维匹配问题 DNA 计算算法 [J]. 电子学报, 2010, 38(8): 1831 – 1836.
Zhou Xu, Li Ken-li, Yue Guang-xue et al. A new DNA computing's algorithm for 3-dimensional matching problem based on divide and conquer strategy [J]. Acta Electronica Sinica, 2010, 38(8): 1831 – 1836. (in Chinese)
- [7] Xiaojun Ma, Lombardi, F., Synthesis of tile sets for DNA self-assembly [J]. Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on, 2008, 27(5): 963 – 967.
- [8] Seeman N C. DNA nanotechnology: Novel DNA constructions [J]. Annual Review of Biophysics and Biomolecular Structure, 1998, 27: 225 – 248.
- [9] Winfree Erik et al. Design and self-assembly of two-dimensional DNA crystals [J]. Nature, 1998, 394: 539 – 544.
- [10] He Yu, Mao Cheng-De et al. Hierarchical self-assembly of DNA into symmetric supramolecular polyhedral [J]. Nature, 2008, 452: 198 – 202.
- [11] Jonoska N, Karl S A, Saito M. Three dimensional DNA structures in computing [J]. Biosystems, 1999, 52: 143 – 153.
- [12] 杨静,张成. DNA 自组装技术的研究进展及难点 [J]. 计算机学报, 2008, 31(12): 2138 – 2148.
Yang Jing, Zhang Cheng. Progress and difficulty in DNA self-assembly technology [J]. Chinese Journal of Computers,

2008, 31(12):2138 – 2148. (in Chinese)

- [13] 高琳, 许进, 张军英. DNA 计算的研究进展与展望[J]. 电子学报, 2001, 29(7):973 – 977.
Gao Lin, Xu Jin, et al. A survey of DNA computing[J]. Acta Electronica Sinica, 2001, 29(7):973 – 977. (in Chinese)
- [14] Bell J, Stevens B. A survey of known results and research areas for N-Queens[J]. Discret. Math. 2009, 309(1):1 – 31.
- [15] H Ahrabian, A Merzaei, A Nowzari-Dalini. A DNA sticker algorithm for solving N-queen problem[J]. International Journal for Computer Science and Applications, 2008, 5(2):12 – 22.
- [16] Tsu-Ju Fu, Nadrian C. Seeman. DNA double- crossover molecules[J]. Biochemistry, 1993, 32:3211 – 3220.
- [17] Brun Y. Solving NP-complete problems in the tile assembly model[J]. Theoretical Computer Science, 2008, 395(1):31 – 46.
- [18] Winfree E. Algorithmic Self-Assembly of DNA[D]. California Institute of Technology. Pasadena, California, 1998.
- [19] E Winfree, The xgrow simulator[EB/OL]. <http://dna.caltech.edu/Xgrow/>, 2009.
- [20] 马润年, 张强, 高琳, 许进. 图的最大权团的 DNA 计算[J]. 电子学报, 2004, 32(1):13 – 16.
Ma Run-nian, Zhang Qiang, et al. Using DNA to solve the maximum weight clique of graphs[J]. Acta Electronica Sinica, 2004, 32(1):13 – 16. (in Chinese)
- [21] Yuriy Brun, Efficient 3-SAT algorithms in the tile assembly model[J]. Natural Computing, 2011, 11(2):209 – 229.
- [22] 殷志祥, 崔建中, 支凌迎, 等. 可满足问题的分子信标计算模型[J]. 计算机学报, 2008, 31(12):2200 – 2206.

Yin Zhi-Xiang, Cui Jian-Zhong, et al. Molecular beacon based DNA computing model for general satisfiability problem[J]. Chinese Journal of Computers, 2008, 31(12):2200 – 2206. (in Chinese)

- [23] S. Garcia, A. Orailoglu. Making DNA self-assembly error-proof: Attaining small growth error rates through embedded information redundancy[A]. Design, Automation & Test in Europe Conference & Exhibition, 2009. DATE'09[C]. Nice, 2009. 898 – 901.

作者简介



吴帆 男, 1973 年生于湖南涟源, 湖南大学信息科学与工程学院讲师, 在职博士生, 研究方向为 DNA 计算和并行计算。
E-mail: wufan_w hu@163.com



李肯立(通信作者) 男, 1971 年生于湖南涟源, 湖南大学信息科学与工程学院教授, 博士生导师. 研究方向为 DNA 计算和并行计算。
E-mail: lk1510@263.net