# 一种求解图着色问题的优化组合遗传算法①

帅训波[1] 杨遂发[1] 周兆华[2] 王建忠[1] (1.中国石油勘探开发研究院 廊坊分院地球物理与信息研究所 河北 廊坊 065007；2.中国石油勘探开发研究院 廊坊分院天然气开发研究所 河北 廊坊 065007)

**摘 要：** 图着色算法是一种典型的NP-完全问题。在逆序算子、对偶算子和矩阵遗传算子的性能研究基础上，采用自然数与二进制相互转换的编码方案，应用图着色问题的约束条件建立适应度评价函数，将具有良好局部搜索性能的矩阵遗传算子与具有良好局部搜索性能的逆序与对偶组合算子优化组合应用，构造了一种用于求解图着色问题的优化组合遗传算法，保证了算法的全局收敛性。与基本遗传算法相比较，实验结果表明，该算法对图着色问题有较好的求解性能。

**关键词：** 图着色；遗传算法；逆序与对偶组合算子；矩阵遗传算子

## Optimization Combination Genetic Algorithm for Graph Coloring Problem

SHUAI Xun-Bo[1], YANG Shui-Fa[1], ZHOU Zhao-Hua[2], WANG Jian-Zhong[1]

(1. Langfang Branch of Research Institute of Petroleum Exploration and Development, Institute of Geophysics and Information, PetroChina, Langfang 065007, China; 2. Langfang Branch of Research Institute of Petroleum Exploration and Development, Institute of Oil Field Development, PetroChina, Langfang 065007, China)

**Abstract:** Graph coloring is a NP-Complete problem. In this paper, based on the research result of Matrix genetic operator, inverse and dual combination operator, an optimization combination genetic algorithm is constructed through the inverse and dual combination genetic operator combined with matrix genetic operator to solve Graph Coloring problem. A transform coding between integer and binary is introduced to make good use of the combination operators. Fitness function based on constraint of Graph Coloring problem is designed, and the convergence of algorithm is proved. Better efficiency of the optimization combination genetic algorithm for solving Graph Coloring problem is verified compared with current genetic algorithm.

**Keywords:** graph coloring; genetic algorithm; inverse and dual combination operator; matrix genetic operator

## 1 引言

图着色问题已被广泛直接应用到存储分配、电路布线、工序及排课表等工业问题中，也是现代图论中重要研究的课题之一。然而，图着色问题是一个NP-完全问题，不存在多项式时间的精确算法，因此，对它的求解算法一直是研究的热点。遗传算法应用于求解图着色问题，取得了不俗的效果[1-3]，但研究表明，一般遗传算法在求解图着色问题时，需要多次产生初始种群，且求解效率和质量均不理想[1,4,5]。针对这一弊端，本文在逆序算子、对偶算子和矩阵遗传算子研究基础之上[6-9]，构造了一种新型优化组合遗传算法，采用自然数与二进制相互转换的编码方案，以图着色问题约束条件为基础，设计适应度函数，保证了算法的全局收敛性。实验结果表明，本文算法对求解

① 收稿时间:2009-11-17;收到修改稿时间:2010-01-28

图着色问题有较好性能。

## 2 问题描述

图的一个正常点着色，就是把图的顶点集合分成个独立集合的一种分类，令表示种颜色，则图的点着色定义：

定义 1. 图的点着色是从图的顶点集到的一个映射，当且仅当且时，。图的所有点着色构成的集合，记为，若，则称图中点可着色的。

定义 2. 使图为点着色的最小颜色数称为图的色数，记为。

## 3 优化组合遗传算法

一般遗传算法在求解图着问题时，对给定的种颜色(令)，产生初始种群，在遗传算子的作用下，求解着色；若找到方案，令，重新产生初始种群，求解着色，直到找到问题的着色方案为止。显然，这种多次产生初始种群的过程，浪费计算时间，且损失种群中已有的分割信息，整体求解效率较低[5]。本文在逆序算子、对偶算子和矩阵遗传算子性能研究基础之上，构造了一种新型优化组合遗传算法，根据图着色问题的约束条件设计适应度评价函数，采用自然数与二进制互换的编码方案，从较小规模的初始种群出发，向着色方案搜索逼近。

### 3.1 新型遗传算子

定义 3 (逆序运算).若染色体的反序列基因组有意义，逆序运算实现由染色体顺序基因组 $s=\ \alpha_1\ \alpha_2\ \dots\ \alpha_{n-1}\ \alpha_n$ 得到其逆序基因组 $s'=\alpha_n\ \alpha_{n-1}\dots\alpha_2\ \alpha_1$。

定义 4 (对偶运算). 对偶运算实现依次对基因组 $s=\alpha_1\ \alpha_2\dots\alpha_{n-1}\ \alpha_n$ 的每个基因位 $j$，将 $j$ 取值 $\alpha$ 被其任意等位基因 $\beta$ 替代，形成新等位基因组 $s'=\beta_1\beta_2\dots\beta_{n-1}\ \beta_n$。

定义 5[8] (逆序与对偶组合算子). 逆序与对偶组合算子对基因组 $S$ 顺序完成如下四步操作：

(1)应用逆序运算产生基因组 $S\_R$，应用对偶运算产生基因组 $S\_A$；

(2)对 $S\_R$ 应用对偶产生基因组 $S\_R\_A$，或者对 $S\_A$ 应用逆序产生基因组 $S\_A\_R$；

(3) 染色体 $S\_N=MAX(S\_R, S\_A,\ S\_R\_A$ 或 $S\_A\_R)$，选取适应度最好的染色体；

(4)若 $S\_N$ 的适应度优于 $S$ 的适应度，则 $S$ 在种群的位置被 $S\_N$ 替代，反之，$S\_N$ 被淘汰。

定义 6 [9] (群体矩阵). 在二进制编码中，编码长度为 $n$，根据给定方法选取 $n$ 个染色体，并将它们随机存储在矩阵 $A_{n\times n}$ 矩阵的各行中，该 $n\times n$ 矩阵 $A$ 称为一个群体矩阵。

定义 7 [9] (矩阵遗传算子).矩阵遗传算子是实现将群体矩阵 $A_{n\times n}$ 转置，得到转置矩阵 $A^T$，以 $A^T$ 每行基因组成新染色体，将 $A^T$ 中 $n$ 个新染色体与 $A$ 中 $n$ 个染色体组成新种群，依据适应度值选取最优的 $n$ 个染色体参与遗传运算。

模式定理分析和大量实验均表明[6-9]，逆序与对偶组合算子对种群表达的样本空间进行“深度”搜索，具有更强的局部搜索能力；矩阵算子向距离搜索目标更近的样本空间“扩展”逼近，具有较强的全局搜索能力。本文将二者组合应用，使构造的遗传算法更加完善。

### 3.2 编码

对给定的图 $G=(V,E)$，其中 $|V|=n$ 且 $|E|=m$，对其采用邻接矩阵方式存储。长度为 $n$ 的数组 $A$ 顺序存储图 $G$ 的顶点，任意给定顶点 $V_i$，则 $A[i]=V_i$；用 $n\times n$ 的二维数组 $B$ 存储图 $G$ 的各边，任意给顶点对 $(V_i,V_j)$，若 $(V_i,V_j)\in E$，则 $B[i,j]=1$，否则 $B[i,j]=0$，那么 $\sum_{i=0}^{i=n}\sum_{j=0}^{j=n}B[i,j]=m$。

给定图 $G$ 的顶点数 $n$，求取自然数 $t$，使得 $2^{t-1}\le n\le 2^t$ 成立，那么，对于数组 $A$ 所存储的每个顶点 $a[i]$ 所着颜色 $i\in C(k)$ 可以用 $t$ 位二进制串表示。

在种群初始化每个染色体时，首先从 $C=\{1,2,..,k,2^t\}$ 颜色集合中选取 $n$ 种不同颜色，分别对图 $G$ 的 $n$ 个顶点着色，满足各顶点着色不同，则染色体自然数编码串为 $C_1C_2C_3...C_n$，对其每位基因用 $t$ 位二进制串表示，最后形成染色体的 $n\bullet t$ 位二进制串，然后对这样的二进制编码初始种群，应用 3.4 节的优化组合遗传算法搜索，向最优着色方案的二进制串逼近。在遗传搜索终止时，对所得到的每个最优染色体，按照从左至右顺序每 $t$ 二进制串为一组译成自然数，表示对应顶点所着颜色，得到 $C=C_1C_2C_3...C_n$，对 $C$ 中所包括的不同自然数个数进行统计，记为 $k$，其中

$1 < k \le n$，即 $x(G) = k$。按照统计时的对应顺序，把 $C$ 中各个自然数与集合 $R = \{1,2,...,k\}$ 中各个自然数对应，转换得到最优染色体 $C$ 的着色方案。例如本文算法对有 5 个顶点的图 $G$ 进行着色的一个染色体译码过程如图 1 所示：

二进制串：　001　101　001　101　111

$\Downarrow$

自然数串：　1　5　1　5　7

$\Downarrow$

统计 $k = 3$，$R = \{1,2,3\}$；　$1 \to 1$，$5 \to 2$，$7 \to 3$

$\Downarrow$

最优着色方案：　1　2　1　2　3

图 1　染色体译码

### 3.3 适应度评价函数

对二进制串染色体评价，首先将其译成自然数编码串 $C = C_1C_2C_3...C_n$，然后遍历存储图 $G$ 的 $n \times n$ 邻接矩阵的上三角矩阵，所访问的边与 3.2 节二维数组 $B$ 中 $B[i,j]$ 相对应，若 $B[i,j] = 1$，则进一步判断顶点 $i$ 和顶点 $j$ 的着色是否满足约束条件，即比较 $C = C_1C_2C_3...C_n$ 中的 $C_i$ 和 $C_j$ 是否相等。因此，根据任意相邻顶点要着不同色的约束条件构造函数：

$$f_1(x) = \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} x_{ij} \qquad (1)$$

$$\text{其中 } x_{ij} = \begin{cases} 1，当 B[i,j] = 1，且 C_i = C_j； \\ 0，其它； \end{cases}$$

通过式(1)对染色体中违反相邻顶点要着不同色的约束条件进行评价度量，$f_1$ 函数值越小，染色体满足此约束条件的性能越好。

给定自然数编码染色体 $C = C_1C_2C_3...C_n$，由于图 $k-$点着色问题定义知，$C$ 中所含颜色数目大于 1，对 $C$ 中所含不同的颜色数目评价构造函数：

$$f_2(y) = 1 + \sum_{i=1}^{n-1}\prod_{j=i+1}^{n} y_{ij} \qquad (2)$$

$$\text{其中 } y_{ij} = \begin{cases} 1，当 C_i \ne C_j； \\ 0，当 C_i = C_j； \end{cases}$$

由式(2)可得，$f_2$ 函数值越小，染色体所满足的颜色数目约束条件越好。

根据图 $k-$点着色问题定义知，满足 $x(G)$ 最优染色体的 $f_1 = 0$，且 $f_2 = k$。综合式(1)和式(2)所表述的两个约束条件度量，构造染色体的最终适应度评价函数为：

$$\min F(x,y) = \max[f_1(x), f_2(y)] \bullet [f_1(x) + f_2(y)] \qquad (3)$$

定理 1．求解图 $G$ 的 $k-$点着色问题，自然数编码染色体 $C = C_1C_2C_3...C_n$，若 $f_2(C) < k$，则染色体 C 必不是图 $G$ 的 $k-$点着色问题的解。

证明：反证法。假设染色体 C 是图 $G$ 的 $k-$点着色问题的一个解，那么由定义 2 知，$f_2(C) = k$，这与已知条件 $f_2(C) < k$ 相矛盾。故定理 1 得证。

定理 2．求解图 $G$ 的 $k-$点着色问题，自然数编码染色体 $C = C_1C_2C_3...C_n$，若 $f_2(C) \ge k$，则 $F(C) \ge k^2$．

证明：分如下三种情况进行分别讨论：

(1)如果给定的染色体 $C = C_1C_2C_3...C_n$ 是 $k-$点着色问题的一个解，那么由定义 2 得 $f_1(C) = 0$ 且 $f_2(C) = k$，分别代入式(3)计算，则 $F(C) = k^2$；

(2) 如果给定的染色体 $C = C_1C_2C_3...C_n$ 不是 $k-$点着色问题的解，且 $f_2(C) > k > 0$；根据定义 1，若 $C = C_1C_2C_3...C_n$ 满足任意相邻顶点要着不同色的约束条件，即 $f_1(C) = 0$，那么 $F(C) > k^2$；

(3) 如果给定的染色体 $C = C_1C_2C_3...C_n$ 不是 $k-$点着色问题的解，且不满足任意相邻顶点要着不同色的约束条件，且 $f_2(C) > k > 0$，$f_1(C) > 0$；当 $f_2(C) \ge f_1(C) > 0$ 时，$\max[f_1(C), f_2(C)] = f_2(C) > k$，$f_1(C) + f_2(C) > f_2(C) > k$，则 $F(C) > k^2$；当 $f_1(C) > f_2(C) > k$ 时，$\max[f_1(C), f_2(C)] = f_1(C) > k$，$f_1(C) + f_2(C) > f_1(C) > k$，则 $F(C) > k^2$。

综合上述三种不同情况，故定理 2 得证。

定理 3．求解图 $G$ 的 $k-$点着色问题，如果自然数编码染色体 $C = C_1C_2C_3...C_n$ 满足 $F(C) = k^2$，那么染色体 $C$ 必是图 $G$ 的 $k-$点着色问题的一个解。

证明：分如下两种情况进行分别讨论：

(1)若 $f_2(C) < k$，由定理 1 知，该自然数编码染色体 $C$ 必不是图 $G$ 的 $k-$点着色问题的解。

(2)若 $f_2(C) \ge k$，用反证法。假设满足 $F(C) = k^2$ 的该染色体 $C$ 不是图 $G$ 的 $k-$点着色问题的解。

由定理 2 中的第(2)种和第(3)种情况的分别讨论，可得 $F(C) > k^2$，这与 $F(C) = k^2$ 条件相矛盾。故染色体 $C$ 必是图 $G$ 的 $k-$点着色问题的一个解。

综合上述两种情况，故定理 3 得证。

由定理 1、定理 2 和定理 3 可知，本文根据图着

色问题约束条件而构造的适应度评价函数是有效的。

### 3.4 算法构造

本文应用逆序与对偶组合算子和矩阵遗传算子，构造优化组合遗传算法(IDM_GA)流程如下：

步骤 1 根据 3.2 节所述初始种群生成法，得到二进制编码初始种群，并应用逆序与对偶组合算子优化；

步骤 2 排序选择出等于染色体编码长度的染色体，应用矩阵遗传算子进行全局搜索；

步骤 3 对矩阵遗传算子搜索得到的 $n$ 个新染色体，分别应用逆序与对偶组合算子进行局部搜索；

步骤 4 应用 3.3 节的适应度评价函数，计算各个染色体适应度，并根据适应度值对种群进行排序；

步骤 5 判断是否满足算法结束条件，如果满足，则对所有的最优染色体输出，否则转向步骤 2；

步骤 6 将各个二进制编码染色体转换成自然数编码染色体，输出 $x(G)$ 。

引理 1[6]. 如果变异概率 $p_m \in (0,1)$ ，交叉概率 $P_c \in (0,1)$ ，同时采用排序选择和局部搜索策略，遗传算法最终收敛到全局最优解。

定理 4.对图 $G$ 进行求解着色问题，顶点数为 $N$ ，当种群规模大于 $2N \bullet \left\lceil \log_2^N \right\rceil$ 时， IDM_GA 算法以概率 1 收敛到全局最优解。

证明：由文献[9]的矩阵遗传算子性能分析，矩阵遗传算子等效于 $n$ 个染色体的“矩阵”方式杂交；根据文献[6-8]对逆序算子和对偶算子的性能分析，逆序与对偶组合算子，等效于对矩阵遗传算子搜索过程中的 $2n$ 个染色体进行“逆序与对偶”方式变异。设图的顶点数为 $N$ ，由本文 3.2 节编码方法可知，染色体的二进制编码长度为 $N \bullet \left\lceil \log_2^N \right\rceil$ 。如果种群规模大于 $2N\left\lceil \log_2^N \right\rceil$ ，使得“矩阵遗传”式杂交概率 $P_c \in (0,1)$ ，且“逆序与对偶”式变异概率 $p_m \in (0,1)$ ，又因为算法采用选择排序策略，故由引理 1 可知， IDM_GA 算法以概率 1 收敛到全局最优解。

## 4 实验结果与分析

为了验证 IDM_GA 算法对图着色问题的求解性能，用 Java 语言编程实现，选取 3 个典型算例对算法性能测试，其中文献[10]中对 7 个顶点图 $G$ 着色作为算例 1，文献[11]的正十二面体和正二十面体所对应的平面图着色分别作为算例 2 和算例 3，并与基本遗传算法(GA)求解结果相比较。保证两种算法对比的公平性，IDM_GA 算法和 GA 算法种群规模均设为 500，遗传终止代数设为 800 代，GA 算法采用均匀杂交概率 $P_c = 0.45$ 、单点变异概率 $p_m = 0.05$ ，在 CPU 为双 Intel 内核 2.0G，内存 1G 的 PC 机上实验。每种算法各时自运行 200 次，随机抽取 10 次的实验结果，求取对各算例搜索到所用最小颜色的平均数、最小颜色有效着色方案(最优解)平均个数、最优解的平均迭代代数及得到最优解时的遗传迭代平均时间，结果如表 1 和表 2 分别所示：

表 1 GA 算法的求解结果

| 算例 | 色数 | 最优解个数 | 代数 | 时间 |
|---|---|---|---|---|
| 算例 1 | 3.8 | 0.6 | 307.6 | 1472.3 |
| 算例 2 | 5.2 | 0.3 | 632.4 | 3128.6 |
| 算例 3 | 5.4 | 0.1 | 685.9 | 3928.1 |

注：时间统计单位为毫秒。

表 2 IDM_GA 算法的求解结果

| 算例 | 色数 | 最优解个数 | 代数 | 时间 |
|---|---|---|---|---|
| 算例 1 | 3 | 3 | 29.5 | 962.6 |
| 算例 2 | 4.4 | 2.2 | 41.3 | 1825.9 |
| 算例 3 | 3.5 | 1.8 | 46.7 | 2316.4 |

注：时间统计单位为毫秒。

从表 1 和表 2 的实验结果对比看出，本文 IDM_GA 算法可以在较少的时间内寻找到较多的最优有效着色方案，与 GA 算法的求解效果比较，IDM_GA 算法对图着色问题有着良好的整体求解性能。此外，IDM_GA 算法运行参数只需保证全局收敛性的种群规模，比 GA 算法更简洁高效。

## 5 结论

对被求解的着色图采用邻接矩阵存储，提出了自然数与二进制互换的编码方案。从图着色问题的约束条件出发，设计了有效的适应度评价函数。在逆序算子、对偶算子和矩阵遗传算子的性能分析研究基础之，构造了一种新型的优化组合遗传算法，通过种群规模控制保证了算法良好全局收敛性，将其用于求解图着

(下转第 48 页)

(上接第 77 页)

色问题，与基本遗传算法对比的实验结果表明，本文的优化组合遗传算法有着较好的求解性能.

## 参考文献


1 Eiben AE. Vender Hauw JK, Van Hemert JI. Graph Coloring with Adaptive Evolutionary Algorithms. Journal of Heuristics,1996,4(1):16－24.

2 Fleurent C, Ferland JA. Genetic and Hybrid Algorithms for Graph Coloring. Annals of Operations Research, 1995,63(3):437－463.

3 Morgenstern C. Distributed Coloration Neighborhood Search.Discrete Mathematic and Theoretical Computer Science,1996,26(5):335－358.

4 曹莉,程灏,许钟.一种应用于图着色问题的新型混合遗传算法.陕西师范大学学报, 2007,35(3):24－27.

5 韩丽霞,王宇平.图着色问题的新遗传算法.西安电子科技大学学报, 2008,35(2):319－313.

6 马书南,帅训波,曹凤雪.一种基于逆序算子的优化组合遗传算法.电子技术应用, 2006,32(6):18－21.

7 帅训波,马书南,周相广.一种基于对偶与逆序组合算子的遗传算法.系统仿真学报, 2009,21(23):7404－7407.

8 帅训波,马书南.一种基于遗传算子优化组合的 TSP 问题求解方法.山东理工大学学报,2009,23(5):29－32.

9 帅训波,马书南,周相广,等.一种基于矩阵遗传算子的优化组合遗传算法.小型微型计算机系统,2009,30(5):951－954.

10 霍红卫,许进,保铮.分组遗传算法用于图的着色.西北民族学院学报, 2000,21(1):5－11.

11 卜月华.图论及其应用.南京:东南大学出版社, 2003.