Find a special connected block 解题报告

绍兴市第一中学 孙耀峰

1 试题来源

Codechef April 2012 CONNECT

2 试题大意

给出一个 $N*M$ 的矩阵。第 i 行第 j 列的格子颜色为 $A_{i,j}$，权值为 $V_{i,j}$。颜色范围为 $[-1, N*M]$。

现在需要在这个矩阵里找出一个连通块。要求颜色为 -1 的格子不能选，且这个联通块至少要包含 K 个不同的正数颜色，要求连通块内格子的权值和最小。

数据范围：$N, M \leq 15, 1 \leq K \leq 7$。

3 算法介绍

3.1 问题简化

我们先考虑这个问题的简化版本。

假设矩阵中只有 K 种不同的正数颜色，这就是非常经典的斯坦纳树问题。

我们记 $F[i][j][k]$ 为当前在矩阵的 (i, j) 点，k 是一个二进制状态，表示联通块中颜色的信息。如果 k 二进制上某一位是 1 则表示当前联通块中已经包含这种颜色，否则说明没有选该颜色。

存在两种转移:

1. $F[i][j][k] + V[x][y] \rightarrow F[x][y][k]$。其中 (x, y) 为与 (i, j) 相邻的格子。

2. $F[i][j][x] + F[i][j][y\ xor\ x] - V[i][j] \to F[i][j][y]$。其中 $x\ and\ y = x$。

初始状态为:如果一个非障碍的格子 (i, j) 有正数颜色 $A_{i,j}$,则 $F[i][j][2^{A[i][j]-1}] = V[i][j]$，否则 $F[i][j][0] = V[i][j]$。

单次斯坦纳树的效率是 $O(2^K NM)$。最后答案就是对于所有非障碍的格子 (i, j)， $F[i][j][2^K - 1]$ 的最小值。

3.2 算法思路

回到原问题，我们来考虑颜色总数大于 K 种的情况。

我们可以将所有的颜色随机给它一个 $[1, K]$ 之间的数，表示把它的颜色重标号。之后对于矩阵中所有的格子，把它的颜色变成新的标号。

因为现在整个矩阵的颜色不超过 K 种了，所有我们可以效仿之前的斯坦纳树做法，求出这种重新标号下的最优解。

容易发现，这种重标号只可能使答案变劣，且肯定都是合法的。实际随机中，因为我们要保证 $[1, K]$ 这些标号至少有矩阵中的一种颜色对应。所以我们需要先从所有颜色中挑出 K 种颜色标号为 $1..K$，之后对于没有选出的颜色随机一个 $[1, K]$ 之间的标号。

3.3 正确性证明

现在我们考虑这种随机方案进行单次的正确性。

我们假设最后最优答案的联通块中含有的 K 种颜色为 $A_1, A_2, .., A_K$。那么只要 $A_1, A_2, .., A_K$ 这 K 种颜色随机到的标号各不相同，也就是说是 $1 - K$ 的一个排列，那么在斯坦纳树中就可以得到最优的答案。而对于另外的颜色，无论它随机到的标号是什么，都对得到最优解没有干扰。

也就是说，单次的正确概率为:

$$\frac{K!K^{T-K}}{K^T} = \frac{K!}{K^K}$$

其中 T 为矩阵中不同的颜色数。

假设我们进行 $Time$ 次随机，这 $Time$ 次随机得到最优答案的概率为:

$$1 - (1 - \frac{K!}{K^K})^{Time}$$

由于本题中 $K \le 7$。经过计算，只需进行 700 次的随机就可以保证正确性达到 99% 了。我提供的标准程序也正是进行 700 次斯坦纳树。

时间复杂度 $O(Time * 2^K NM)$，空间复杂度$O(2^K NM)$。