WinterCamp 2003 论文　芜湖一中　许智磊

浅谈补集转化思想
在统计问题中的应用

前言

统计问题，是我们经常遇到的一类问题

通常认为统计问题是对满足某些性质的对象进行计数的问题

其解法或多或少地建立于**枚举**之上

"枚举"往往是低效的代名词！！

前言

很多时候，我们就需要一些技巧来降低统计的时间复杂度

离散化和**极大化**思想、**二分法**、**事件表**等方法经常可以起到很好的效果。

因此它们作为常规的统计方法，在解题时首先被想到。

前言

然而这些常规方法也有不能奏效的时候

这时我们就需要一些非常规的方法来解决问题

其中的一种就是利用补集转化思想来帮助解决统计问题

补集转化思想在很多方面有着广泛的应用，让我们来看看在解决统计问题方面它又有哪些精彩表现吧！

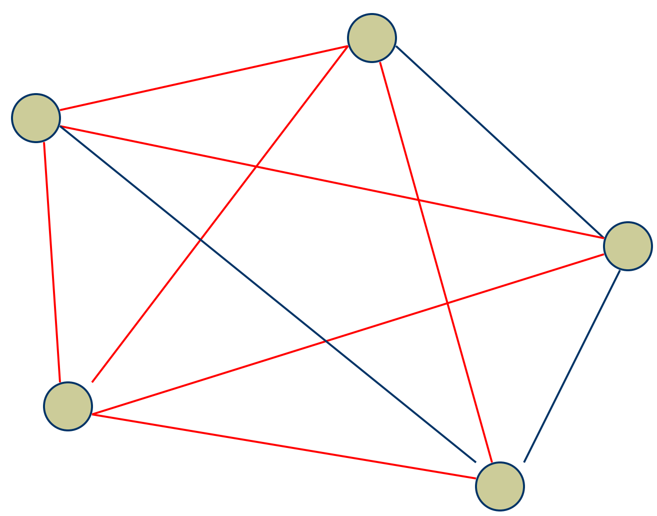
例一　单色三角形问题（ POI9714 TRO ）

题目大意

输入点数 n 、红色边数 m 以及这 m 条红色的边所连接的顶点标号，输出单色三角形个数 R 。

$3<=n<=1000$ ， $0<=m<=250000$ 。

用形。



初步分析

自然的想法：用一个数组记录每两点间边的颜色。枚举所有的三角形（这是通过枚举三个顶点实现的），判断它的三边是否同色，若同色则总数 R 加 1 （当然，初始时 R 为 0 ）。

空间上：　$O(n^2)$，需要一个 1000*1000 的大数组
时间上：　$O(n^3)$， n 达到 1000， 无法接受！
常用技巧：无从下手。

深入思考

本题中单色三角形的个数可以非常庞大，所以一切需要枚举每个单色三角形的方法都是不可能高效的。

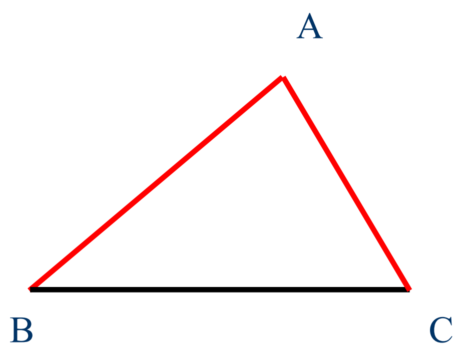
单纯的枚举不可以，那么组合计数是否可行呢？

从总体上进行组合计数很难想到。我们尝试枚举每一个点，设法找到一个组合公式来计算以这个点为顶点的单色三角形的个数。

深入思考

组合公式很难找到！

原因：从一个顶点 A 出发的两条同色的边 AB 、 AC 并不能确定一个单色三角形 ABC ，因为 BC 边有可能不同色。



边 $\not\longrightarrow$ 单色三角形

补集转化

从反面来看问题：每两点都有边连接，所以每三个点都可以组成一个三角形（单色或非单色的），所有的三角形数 S=C(n,3)=n*(n-1)*(n-2)/6 。

单色三角形数 R 加上非单色三角形数 T 就等于 S，所以如果我们可以求出 T，那么显然， R=S－T 。

原问题转化为：怎样高效地求出 T

补集转化

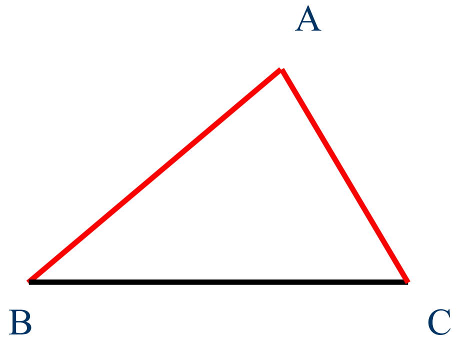
原先的枚举＋组合计数算法的障碍是无法在“边”与“单色三角形”之间建立确定的对应关系。

边 $\xrightarrow{\text{YES!!}}$ 非单色三角形

补集转化

非单色三角形的三条边共有红黑两种颜色

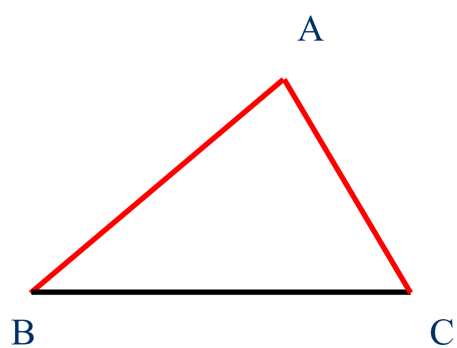
其中两条边同色，另一条边异色



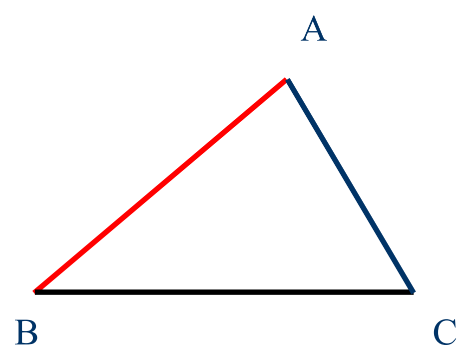
一个非单色三角形 ⟶ 两对“有公共顶点的异色边”

补集转化

如果从一个顶点 B 引出两条异色的边 BA 、 BC ，则无论 AC 边是何种颜色，三角形 ABC 都只能是一个非单色三角形



OR



一对“有公共顶点的异色边” $\longrightarrow$ 一个非单色三角形

补集转化

非单色三角形数 T=“有公共顶点的异色边”的总对数 Q/2

Q 很容易求出：

每个顶点有 n-1 条边，根据输入的信息可以知道每个顶点 i 的红边数 E[i]，那么其黑边数就是 n-1-E[i]。根据乘法原理，以 i 为公共顶点的异色边的对数就是 E[i]*(n-1-E[i])，所以

$$Q = \sum_{i=1}^{n} E[i]*(n-1-E[i])$$

补集转化

Q 求出之后， $R=S-T=n*(n-1)*(n-2)/6-Q/2$

时间复杂度：$O(m+n)$

空间复杂度： $O(n)$

优秀的算法！

小结

通过补集转化，我们在原来无法联系起来的“边”和“三角形”之间建立起确定的关系，并以此构造出组合计数的公式。

单纯的枚举 ⟷ 枚举＋组合计数

这个例子中补集转化思想的作用：

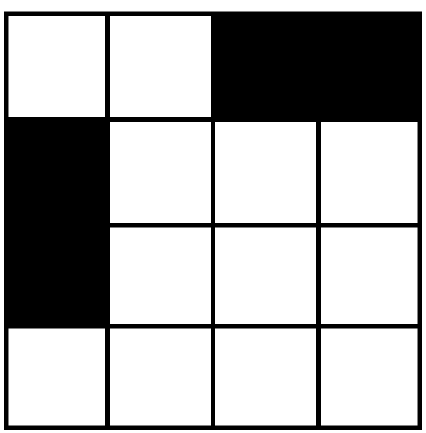
为找到一个本质上不同的算法创造了条件

例二　海战游戏（改编自 Ural1212 Sea Battle ）

题目大意

海战游戏是在一个 N 行 M 列的方格棋盘上摆放“军舰”，一艘军舰是连在一起的 X 行 Y 列方格，每个方格都全等于棋盘上的格子，于是军舰就可以摆放在棋盘上，使军舰的每个格子和棋盘的格子重合。

摆放时必须遵守如下规则：任意两艘军舰的任意两个格子不得重合或在八个方向（即上、下、左、右、左上、右上、右下、左下）上相邻。



例如左图中已经摆放了一个 1 行 2 列的军舰和一个 2 行 1 列的军舰，如果我们要再摆一个 1 行 2 列的军舰，有两种方案。如果要再摆一个 2 行 2 列的军舰，只有一种方案。

初步分析

枚举每一种摆放方案并判断是否符合规则显然不可接受

原题就是给定了一个网格，上面某些矩形区域已经被占用，现在要在里面放入一个新的矩形，不能和已被占用的格子重合或是相邻。

这是典型的在有障碍点的网格上求摆放方案数的统计问题。

看起来如此经典的问题，用常规的方法能否解决呢？

初步分析

实际上，用离散化可以设计出能够接受的算法。

离散化的算法时间复杂度为 $O(\min\{M,N\}*L)$

虽然对于原题勉强可以应付，但是一旦数据规模再稍稍扩大一点，必定超时。

而且离散化的算法思考比较复杂，编程比较烦琐。

几个工具

在进一步地思考之前，我们先明确几个小问题，以作为下面研究的工具

在一个 X 行 Y 列的矩形 A 中放入一个 P 行 Q 列的矩形 B，共有多少种摆放方案?

结论一：

矩形 B 能够放入矩形 A 中的充要条件是 X>=P 且 Y>=Q，所以如果 X<P 或 Y<Q，方案数为 0 。

否则矩形 B 的左上角可以位于矩形 A 的 1 至 X-P+1 行，1 至 Y-Q+1 列，也就是总共有 (X-P+1)*(Y-Q+1) 种摆放方案。

几个工具

矩形 A 的左上角为 (AX1,AY1)，右下角为 (AX2,AY2)，矩形 B 的左上角为 (BX1,BY1)，右下角为 (BX2,BY2)，如果存在某两个格子 a∈A，b∈B 且 a、b 相邻或重合，就称 A 和 B“相交”。如何判断 A、B 是否相交？

这个问题稍稍复杂一点，但是仔细分析各种情况之后可以得出

结论二： A 和 B 相交的充要条件是

(AX1<=BX2+1) and (AX2>=BX1-1)

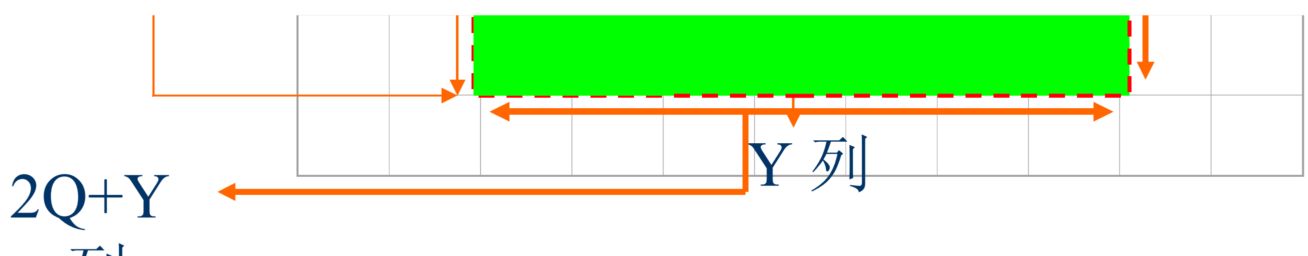
and (AY1<=BY2+1) and (AY2>=AY1-1)。

几个工具

如动画所示，正中的黑色矩形代表已摆放的矩形 A，闪烁的绿色矩形代表新矩形 B 能够与 A 相交的所有方案。

设一个已经摆放的矩形 A 为 X 行 Y 列，新摆放的矩形 B 为 P 行 Q 列，矩形 B 怎样摆放才能和矩形 A 相交呢？根据结论二我们直接就可以得出

结论三：矩形 B 能够与 A 相交的所有摆放方案位于一个 2P+X 行，2Q+Y 列的矩形框内。这个矩形框是在矩形 A 的上、下各扩展 P 行，左、右各扩展 Q 列得到的。



几个工具

所以我们把**结论三**改为：

矩形 B 能够与 A 相交的所有方案位于一个**最多** 2P+X 行，**最多** 2Q+Y 列的矩形框内。

补集转化

符合规则的摆放方案数 R ＝

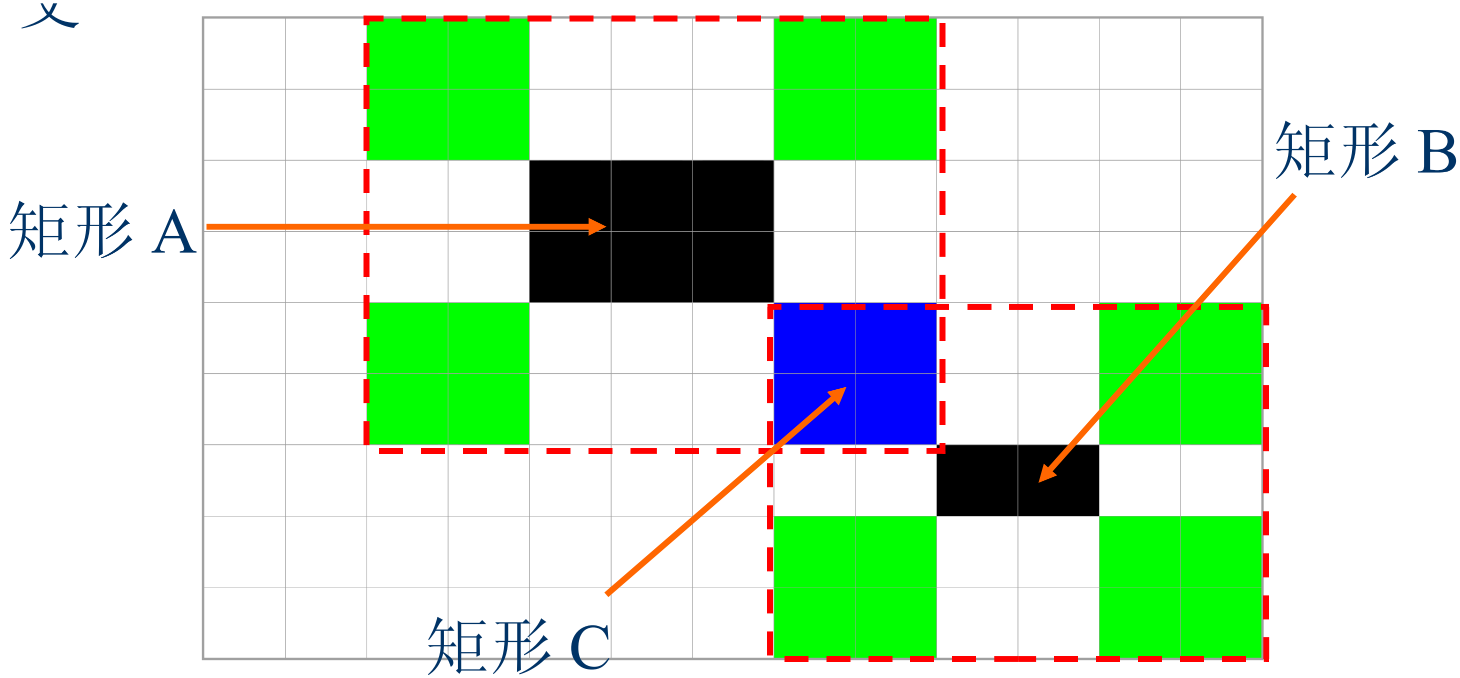
总共的摆放方案数 S －违反规则的摆放方案数 T
＝总共的摆放方案数 S
根据结论一， S 可以根据公式计算出来

问题转化为：怎样高效地求出 T

T 也就是所有与已经摆放的军舰相交的方案数

补集转化

例如我们规定矩形 A 编号较小，则在处理矩形 B 时，方案 C 就不允许计入总数，因为矩形 B 并不是与方案 C 相交的编号最小的已摆放矩形。



补集转化

为了做到这一点，我们只需采取如下算法：

依次处理每个已经摆放的矩形，设当前处理的矩形编号为 i 。

在这个矩形周围一一枚举与它相交的摆放方案。

对于每个方案，再依次枚举编号为 1,2,……(i-1) 的矩形，判断这些矩形能否与当前枚举的方案相交，如果发现有相交的情况，则此方案不能计入总数 T ，否则就将 T 加 1 。

补集转化

根据结论三，与每个已摆放的 X 行 Y 列的矩形相交的摆放方案位于它周围的一个矩形框内，这个矩形框最多 2P+X 行，最多 2Q+Y 列。

再根据结论一，在其中摆放 P 行 Q 列的矩形最多只有 (P+X+1)*(Q+Y+1) 种方案 。

由于每个矩形的大小均在 P*Q 这样的级别，所以总共需要处理的方案数规模为 O(P*Q*L) 。

补集转化

总共需要处理的方案数规模为 $O(P*Q*L)$

处理每个方案的复杂度为 $O(L)$

整个算法的复杂度仅为

$O(P*Q*L^2)$

补集转化

在时间复杂度上，大大领先于离散化的常规解法

思维复杂度、编程复杂度较低

小结

本题从正面考虑，枚举量太大，所以常规的解法是采用离散化技巧来减少枚举量。

但是从反面考虑，枚举量就非常小了。补集转化思想在这里起到的作用是

帮助我们选择了合适的枚举对象，从而减少了枚举量。

总结

比较：两个例子都是利用补集转化思想解决统计问题

相同点　　　　　　　求解目标 R 较困难

求 R 的补集 T 以及 R 、 T 的总和 S 相对较容易

不同点	作用效果	意义价值
例一中	指导我们设计出了本质不同的新算法	似乎是解决这个问题的唯一可行方法
例二中	只是通过改变枚举对象减少了枚举量	比常规方法更自然、更优秀的另一种方法

总结

补集转化思想应用于统计问题的形式是多种多样的，可能从解决问题的各个方面帮助我们。

补集转化思想不仅可以应用于一些非常规的统计问题，而且对于一些常规算法能够解决的问题，应用补集转化思想也许可以做得更好。

总结

补集转化思想，体现了**矛盾对立统一，互相转化**的一种哲学观念。在统计问题中灵活地应用补集转化思想，往往可以起到“出奇制胜”的效果，而这就要求我们注意**培养逆向思维的能力**，才能用好、用活补集转化思想。

值得注意的是，利用补集转化思想解决统计问题作为一种非常规的统计方法，和一些常规的统计方法、技巧之间的关系是辨证的。虽然在本文的例子中，补集转化思想都优于常规方法，但是并不能认为常规方法一定不如非常规方法。大多数的统计问题，还是适合使用常规方法的。

总结

只有将常规方法和非常规方法都灵活地掌握，并对于具体问题选择合适的方法，才能够游刃有余地解决统计问题。

感谢

我的演讲到此结束，谢谢大家！

Thank you all！！！