浅谈容斥原理

成都七中 王迪

各位老师、同学，尊敬的评委们，大家下午好。我是来自成都七中的王迪。我很荣幸能够站在这里跟大家分享一些我的学习经验，而今天我想要讲的主题是容斥原理。

相信大家在小学时就碰到过这样的数学问题：某班有 a 个人擅长唱歌，b 个人擅长画画，c 个人既擅长唱歌也擅长画画，问有多少人有至少一种特长？

我们可以用各种各样的方法来解决这个问题。比如，我们可以利用文氏图，在一张纸上画两个圆，圆 A 表示擅长唱歌的同学的集合，圆 B 表示表示擅长画画的同学的集合，那么它们的相交区域 C 就表示既擅长唱歌也擅长画画的同学，由图我们发现，若我们直接将 A 的人数和 B 的人数相加，会导致 C 中的每个人都被多算一次，所以要减去 C 的人数，即 $a+b-c$。

这个问题中我们使用的方法其实就是容斥原理。在论文中，我从这个最基本的问题出发，通过一些例题讲解了容斥原理在组合计数中的应用，总结了用解决此类问题的思想方法，然后讨论了容斥原理在其他领域内的推广，最后研究了容斥原理更加一般的形式。遗憾今天时间有限，我在这里选择了我的一个原创题与大家分享。当然，我们先来看看容斥原理到底是什么。

类比最开始的那个问题，所有同学组成了一个有限的全集 U，擅长唱歌称为性质 P_1，具有这个性质的同学组成集合 S_1，擅长画画称为性质 P_2，具有这个性质的同学组成集合 S_2，那么我们可以用数学语言来描述原问题，即，S_1 并 S_2 表示至少有一项特长的同学组成的集合，S_1 交 S_2 表示同时有两项特长的同学组成的集合，那么我们所求的就是集合 S_1 并 S_2 的大小，根据之前的答案，我们知道它等于 S_1 的大小加上 S_2 的大小再减去 S_1 交 S_2 的大小，如屏幕所示。

我们将问题稍加推广，当我们用三种性质描述元素时，仍然用画文氏图的方法分析，可以得到下面的结论。我们发现，整个表达式是加减交错的，而且将求并集的问题转化成求交集，于是我们可以归纳出容斥原理的形式。

设我们用 n 种性质 P_1, P_2 到 P_n 描述集合 U 中的元素，满足性质 P_i 的元

素组成了集合 S_i，那么我们可以得出容斥原理的表达式，如屏幕所示。通俗的说，就是满足一个性质的减去满足两个性质的加上满足三个性质的，依次类推。这个式子的证明，我简单提一下，我们可以对每个元素，计算式子两边对它统计的次数，过程中利用组合数进行化简。

让我们来分析一下这个式子。式子的左边是统计若干集合的并，即至少满足一种性质的元素个数；而式子的右边被分成了若干个部分，每个部分是在对特定的集合的交集计数，这提示了我们使用容斥原理解题的一般思路：我们从集合的角度分析问题，找出全集 U 和若干用来描述元素的性质 P_i，然后发现问题是求一些集合的并集大小，我们可以利用容斥原理将其转化，分解为若干的子问题并分别求解，最后合并答案。

关于容斥原理的复杂度，一般来说，因为要枚举子集，对于集合个数 n 是指数级的，但是，如果我们发现，集合的交集的大小只与集合个数有关，就可以枚举集合的个数，利用组合数来计算答案了。

下面我们来看一道比较另类的题目：游戏。

Alice 和 Bob 在玩游戏。他们有一个 n 个点的无向完全图，设所有的边组成了集合 E，于是他们想取遍 E 的所有非空子集，对某个边集 S 有一个估价 $f(S)$，这个估价是这样计算的：考虑 n 个点与 S 中的边组成的图，我们用 m 种颜色对所有点染色，其中同一个联通块的点必须染成一种颜色，那么 $f(S)$ 等于这个图的染色方案数。同时，Alice 喜欢奇数，所以当 S 的大小为奇数时，Alice 的分值加上 $f(S)$，否则 Bob 的分值加上 $f(S)$。求最后 Alice 的分值减去 Bob 的分值的值模 10^9+7 的结果。数据规模 $n, m \leq 10^6$。

我们来看一些最直接的想法。一方面，若按照题目模拟，则需要枚举边集 E 的子集，复杂度太高不能承受；另一方面，若枚举边数，很容易发现，即使两个边集中的边数一样，它们所形成的联通块数不一定一样。我们发现这个问题很难入手，这时我们应该写出问题的数学形式，看能否发现熟悉的东西。

对题意稍加分析，我们可以发现，“同一个联通块的点必须染成一种颜色”等价于“一条边连接的两个点必须染成一种颜色”，这一点非常重要。于是，我们可以设一些变量。设 x_i 表示第 i 个点的颜色，那么一条边 (i,j) 就表示了一个条件 $[x_i = x_j]$。令 $F(C)$ 表示条件 C 下，这个图的染色方案数，其中 C 表示一系列 $[x_i = x_j]$ 的集合。同时我们记 Alice 的得分为 $scoreA$，Bob 的得分为 $scoreB$，那么我们可以直接按照题意来写表达式，如屏幕所示，我们对边

集 E 的所有子集求和，把染色方案用提到的 F 函数代替。而我们的目标是求 $scoreA - scoreB$，我们不妨将两式做差，容易发现奇偶性的条件，因为正负号的原因，可以用 -1 的幂次来表示，然后得到了答案的表达式。我们发现，整个式子的难点在于后面的 F 函数上。

我们考虑 F 函数的含义，对一个边集的子集 S，它代表满足 S 中的所有的边对应的条件的染色方案数。我们再次细化条件，令 T_e 表示满足 e 这条边对应的条件时，所有染色方案组成的集合。容易知道，$F(S)$ 就代表的是一些 T_e 的交集的大小。

那么我们可以把所有的边从 1 到 Q 标号，枚举集合中的边数，重写答案的表达式，如屏幕所示。我们发现上式的右边与容斥原理的形式相同，于是可以尝试逆向运用容斥原理，从式子的右边推到左边，可以发现我们的答案其实是求一些集合的并集的大小，然后我们再来考虑这个并集的含义，是我们在统计一个图 n 个点的染色方案数，使得至少一条边对应的条件成立，即至少两个点的颜色相同。

我再重新描述一下我们得到的新问题：n 个点，m 种颜色，至少两个点的颜色相同，图的染色方案数。我们通过一步简单的补集转化，先求两两颜色不同的的方案数，再用总的方案数去减，这一步就非常简单了，如屏幕所示。

回顾我们的解题过程。首先，我们发现直接入手比较困难，所以选择写出问题的数学形式，用一些设变量或者换元的小技巧进行化简；然后，我们逆向使用了容斥原理，这种思路比较新颖，较难想到；最后，我们发现了问题的本质，算出了答案。

在游戏这道题目，我们以一种不同的方式运用了容斥原理。从这个问题中我发现容斥原理背后蕴藏着一些思想：解决问题时，我们既可以用“隔离法”，拆分条件，解决子问题，再合并答案；也可以用“整体法”，对所求的式子整体感知，逆向思考，合并已知条件，找出问题的本质。这里，体现的都是转化的思想。

另外，在论文中，我还探究了容斥原理在数论和概率论中的推广，这告诉我们不同领域间存在或多或少的联系，在学习过程中，要多联想，多思考，体会所学算法背后蕴含的思想，探寻其本质，形成自己的知识网络，这对我们是大有裨益的。

最后，在这里我想感谢父母对我的养育，感谢我的教练成都七中的张君亮老师以及其他所以给予我帮助的老师，感谢罗雨屏、李凌霄、钟皓曦等同学的帮助，感谢 CCF 给我一个展示自己的机会，最后感谢在座各位的认真听讲。谢谢你们！

我的演讲结束，欢迎提问。