

Combination

组合

问题:

求拥有 n 个元素的集合 $A = \{a_0, a_1, a_2, \dots, a_{n-1}\}$ 中任意取 m 个元素 ($m \leq n$, m 和 n 都是自然数) 的所有组合。

解法:

本文末尾列了很多关于组合算法的文献。本文介绍一种简单易记的算法。从 5 个数字的集合 $A = \{a_0, a_1, a_2, a_3, a_4\}$ 中取出 3 个作为新的集合, 设置排列 $s = [1, 1, 1, 0, 0]$ 表示对这 5 个数字的选择, 第 i 个数字 $s_i = 1$ 表示选择数字 A_i , $s_i = 0$ 表示不选择数字 A_i 。 s 的全排列即为集合 A 的所有组合的映射:

$[1, 1, 1, 0, 0] \rightarrow \{a_0, a_1, a_2\}$
 $[1, 1, 0, 1, 0] \rightarrow \{a_0, a_1, a_3\}$
 $[1, 0, 1, 1, 0] \rightarrow \{a_0, a_2, a_3\}$
 $[0, 1, 1, 1, 0] \rightarrow \{a_1, a_2, a_3\}$
 $[1, 1, 0, 0, 1] \rightarrow \{a_0, a_1, a_4\}$
 $[1, 0, 1, 0, 1] \rightarrow \{a_0, a_2, a_4\}$
 $[0, 1, 1, 0, 1] \rightarrow \{a_1, a_2, a_4\}$
 $[1, 0, 0, 1, 1] \rightarrow \{a_0, a_3, a_4\}$
 $[0, 1, 0, 1, 1] \rightarrow \{a_1, a_3, a_4\}$
 $[0, 0, 1, 1, 1] \rightarrow \{a_2, a_3, a_4\}$

与<Full Permutation>不同的是, s 的全排列中存在重复的情况, 无法照搬。将排列 s 初始化为 $[1, 1, 1, 0, 0]$, 在初始化状态中, 从左开始找到第一组序对 $[1, 0]$ (前 1 个元素为 1, 后 1 个元素为 0), 将这个序对交换位置, 得到 $[1, 1, 0, 1, 0]$, 然后再将序对之前的所有 1 放到排列的前面, 所有 0 放到排列后面, 但都不超过序对本身。

例如对于 $[0, 1, 1, 0, 1]$, 将从左开始的第一组序对 $[1, 0]$ (即 $\{a_2, a_3\}$) 交换后得到 $[0, 1, 0, 1, 1]$, 这时将数组中下标在 $[0, 1]$ 范围中的所有 1 放到前面, 0 放到后面, 得到 $[1, 0, 0, 1, 1]$ 。之后重复的进行该操作, 可以得到:

$[0, 1, 0, 1, 1]$
 $[0, 0, 1, 1, 1]$

直到将所有的 1 都移动到数组的末尾, 再也没有任何序对 $[1, 0]$, 算法结束。根据 s 的全排列生成集合 A 的所有组合即可。该算法时间复杂度为 $C_m^n = \frac{n!}{m!(n-m)!}$ 。

StackOverflow 上关于组合产生算法的问题:

<http://stackoverflow.com/questions/127704/algorithm-to-return-all-combinations-of-k-elements-from-n>

二项式系数:

https://en.wikipedia.org/wiki/Binomial_coefficient

Chase's Twiddle - Algorithm 382: Combinations of M out of N Objects:

<http://dl.acm.org/citation.cfm?id=362502>

<http://www.netlib.no/netlib/toms/382>

Buckles - Algorithm 515: Generation of a Vector from the Lexicographical Index:

<http://dl.acm.org/citation.cfm?id=355739>

https://www.researchgate.net/profile/Bill_Buckles/publication/220492658_Algorithm_515_Generation_of_a_Vector_from_the_Lexicographical_Index_G6/links/5716d7ad08ae497c1a5706ec.pdf

Remark on algorithm 515: Generation of a vector from the lexicographical index combinations:

<http://dl.acm.org/citation.cfm?id=1236470>