

编程之美解题报告

姚大海 (acprimer.yao@gmail.com)

<https://github.com/acprimer/Beauty-of-Programming>

最后更新 2014-12-23

版权声明

本作品采用“Creative Commons 署名 -非商业性使用 -相同方式共享 3.0 Unported 许可协议 (cc by-nc-sa)”进行许可。<http://creativecommons.org/licenses/by-nc-sa/3.0/>

内容简介

编程之美解题报告

数学符号

符号	意义
$lg(x)$	以 2 为底的对数 $\log_2(x)$
$\log_b(x)$	以 b 为底的对数 $\log_b(x)$
$\lfloor x \rfloor$	下取整函数
$\lceil x \rceil$	上取整函数

GitHub 地址

本书是开源的，GitHub 地址：<https://github.com/acprimer/Beauty-of-Programming>

目录

第 1 章 游戏之乐	1	2.3 寻找发帖“水王”	6
第 2 章 数学之魅	2	2.4 1 的数目	8
2.1 求二进制数中 1 的个数	2	2.5 寻找最大的 K 个数	11
2.2 不要被阶乘吓倒	4	2.6 精确表达浮点数	11

第 1 章 游戏之乐

暂无

第 2 章

数学之魅

2.1 求二进制数中 1 的个数

问题

对于一个字节（8bit）的无符号整型变量，求其二进制表示中“1”的个数，要求算法的执行效率尽可能高。

解法一

利用模运算求取 x 的每个二进制位上是否为 1。

```
// 时间复杂度  $O(\lg(x))$ ，空间复杂度  $O(1)$ 
int count_1(int x) {
    int ans = 0;
    while (x != 0) {
        ans += (x % 2);
        x /= 2;
    }
    return ans;
}
```

解法二

利用位运算求取 x 的每个二进制位上是否为 1。位运算比模运算效率要高。

```
// 时间复杂度  $O(\lg(x))$ ，空间复杂度  $O(1)$ 。
int count_2(int x) {
    int ans = 0;
    while (x != 0) {
        ans += (x & 0x01);
        x >>= 1;
    }
    return ans;
}
```

解法三

利用按位与 $x \& (x-1)$ 每次消除一个 1，直到 x 为 0。

$$\begin{array}{rcl}
 1011000 & x \\
 \& 1010111 & x-1 \\
 \hline
 1010000 & x \& (x-1)
 \end{array}$$

图 2-1 $x \& (x-1)$

```

// 时间复杂度 O(M)，空间复杂度 O(1)，其中 M 为二进制中 1 的个数。
int count_3(int x) {
    int ans = 0;
    while (x != 0) {
        x &= (x - 1);
        ans++;
    }
    return ans;
}

```

扩展问题 1

如果变量是 32 位的 DWORD，你会使用上述的哪一种算法，或者改进哪一个算法？

解法

DWORD 是一个 32 位无符号整数，因此该问题与原问题唯一的区别就是数据的范围不同。上述的三种解法都适用，而且都不要做任何改进。另外对于原书中的解法四使用分支操作和解法五查表法是不适用的。时间复杂度和空间复杂度和原解法相同。

扩展问题 2

给定两个正整数（二进制形式表示）A 和 B，问把 A 变为 B 需要改变多少位（bit）？也就是说，整数 A 和 B 的二进制表示中有多少位是不同的？

解法

拿到这个问题，我们首先想到的就是按照从低到高依次检查 A 和 B 的每一位是否相同，然后记录个数，时间复杂度是 $O(\log_2(\text{Max}(A, B)))$ 。接着，我们会想能不能把复杂度降到只和位不同的个数相关。自然的就会想到通过一种位运算把不同的 1 取出来，这种位

运算就是按位异或 \wedge 。计算 $C=A\wedge B$ ，再计算 C 中 1 的个数就行了，可以采用原题中的解法三。

```
// 时间复杂度  $O(M)$ ，空间复杂度  $O(1)$ ，其中  $M$  为  $A$  和  $B$  二进制中不同的个数。
int exercise_2(int A, int B) {
    return count(A ^ B);          // 调用上面的 count 函数，计算  $A\wedge B$  中 1 的个数。
}
```

扩展问题 3

给定整数 N ，判断它是否为 2 的幂，比如 0, 2, 4, 8, ... 都是 2 的幂。这是原书中 2.2 的扩展问题，我感觉放在这里比较合适。

解法

```
boolean isPowerOf2(int N) {
    if (N >= 0 && ((N & (N - 1)) == 0)) return true;
    return false;
}
```

2.2 不要被阶乘吓倒

问题

1. 给定一个整数 N ，那么 N 的阶乘 $N!$ 末尾有多少个 0 呢？例如， $N = 10$ ， $N! = 3628800$ ， $N!$ 的末尾有两个 0。
2. 求 $N!$ 的二进制表示中最低位 1 的位置。

问题 1 解法

对 $N!$ 进行质因数分解得到 $N! = \prod p_i^{k_i}$ ，由于 $10=2\times 5$ ，所以 $N!$ 末尾 0 的个数只和 5 的指数有关。计算 $N!$ 中有多少个 5，可以利用公式 $\lfloor \frac{N}{5} \rfloor + \lfloor \frac{N}{5^2} \rfloor + \lfloor \frac{N}{5^3} \rfloor + \dots$ 。

```
// 时间复杂度  $O(\log_5(N))$ ，空间复杂度  $O(1)$ 
int count_1(int N) {
    int ans = 0;
    while (N != 0) {
        N /= 5;
        ans += N;
    }
    return ans;
}
```

问题 2 解法一

和问题 1 基本相同，只是现在二进制表示中末尾 0 的个数是和 $N!$ 中 2 的指数有关。利用公式 $\lfloor \frac{N}{2} \rfloor + \lfloor \frac{N}{2^2} \rfloor + \lfloor \frac{N}{2^3} \rfloor + \dots$ 。

```
// 时间复杂度  $O(\lg(N))$ ，空间复杂度  $O(1)$ 
int count_base_2(int N) {
    int ans = 0;
    while (N != 0) {
        N >>= 1;
        ans += N;
    }
    return ans;
}
```

问题 2 解法二

$N!$ 中质因数 2 的个数等于 $N - N$ 的二进制表示中 1 的个数。时间复杂度 $O(M)$ ， M 是 N 二进制表示中 1 的个数。这个解法技巧性太强了，不看解答完全想不到，而且扩展性差。

代码略。

扩展问题

求整数 N 的 B 进制表示中末尾 0 的个数。题目来源 UVA 10061。

解法

对 $N!$ 进行质因数分解得到 $N! = \prod p_i^{k_i}$ ，对 B 进行质因数分解得到 $B = \prod p_i^{t_i}$ ，我们用 $f(N, B)$ 来表示 N 的 B 进制表示中末尾 0 的个数，可以得出下面的等式： $f(N, B) = \min\{\frac{k_i}{t_i}\}$ 。

```
// 时间复杂度  $O(\lg_B(N))$ ，空间复杂度  $O(1)$ 
int countZeros(int n, int b) {
    int ans = Integer.MAX_VALUE;
    for (int i = 2; i <= b; i++) {
        if (b % i != 0) continue;
        int cnt = 0;
        while (b % i == 0) {
            b /= i;
            cnt++;
        }
        int tmp = 0, tn = n;
        while (tn != 0) {
            tn /= i;
            tmp += tn;
        }
    }
}
```

```

    }
    ans = Math.min(ans, tmp / cnt);
}
return ans;
}

```

这里我们可以考虑一个问题，可不可以直接求出最大的 $p_i^{t_i}$ ，然后计算 $N!$ 中有多少个 p_i 。
提示：N=4, B=40.

2.3 寻找发帖“水王”

问题

给定一个整型数组，每个数组元素表示一个 ID，数组长度为 N，数组中有一个 ID 出现的次数超过了数组长度的一半，求出这个 ID。

解法 1

最简单的解法就是暴力枚举每一个数组元素，判断该元素是否超过一半。时间复杂度是 $O(N^2)$ 。

然后我们考虑对数组进行排序（很多问题一旦排序之后就会有思路了），这时候数组的中位数一定就是我们要求的 ID。时间复杂度是 $O(N\lg(N))$ 。

最后，我们考虑这样一个场景，让不同的 ID 进行 PK，最后剩下的一定就是超过一半的 ID。可以用一个栈来模拟数组中的元素 PK 的过程。

```

// 时间复杂度 O(N)，空间复杂度 O(N)
int findMostWithStack(int[] ids) {
    Stack<Integer> stack = new Stack<Integer>();
    for(int i=0;i<ids.length;i++) {
        if(stack.isEmpty()) {
            stack.push(ids[i]);
        } else {
            if(stack.peek()==ids[i]) stack.push(ids[i]);
            else stack.pop();
        }
    }
    return stack.peek();
}

```

解法 2

下面我们就来考虑一下是否可以降低空间复杂度。其实上面的做法中我们对于栈中的元素并不关心，我们只关心当前栈顶元素和栈的大小，所以我们就可以不用栈来处理了，用两个变量记录栈的大小和栈顶元素即可。


```
// 时间复杂度 O(N), 空间复杂度 O(1)
int findMost(int[] ids) {
    int candidate = -1, count = 0;
    for (int i = 0; i < ids.length; i++) {
        if (count == 0) {
            candidate = ids[i];
            count = 1;
        } else {
            if (ids[i] == candidate) count++;
            else count--;
        }
    }
    return candidate;
}
```

解法 3

利用哈希表。

扩展问题 1

原问题中保证一定存在一个元素超过一半, 那么如果去掉这个条件呢? 求出这个元素的下标 (数组下标从 0 开始), 如果不存在则返回 -1.

解法

如果不存在超过一半的元素, 上面的两个解法其实并不能正确返回。比如 [1,2,3], 返回的就是 3. 其实这里要做的就是对这个得到的结果进行判断出现次数是否超过一半, 时间和空间复杂度都不改变。

代码略。

扩展问题 2

有 3 个元素出现的次数超过了数组长度的 $\frac{1}{4}$, 找出这 3 个元素。

解法

先排序再查找是可以找出的, 时间复杂度是 $O(N\lg(N))$, 这里就不给出具体代码了。我们同样也可以采用 PK 的方式来决出次数最多的 3 个数字, 这样就需要保存 3 个候选元素和它们对应的次数, 在遇到新的元素时, 先跟这个 3 个候选元素比较, 如果次数为 0 就把当前元素作为候选元素, 如果当前元素与候选元素相等那么把它的次数加 1, 如果当前

元素与 3 个候选元素都不相等就把 3 个候选元素的次数都减去 1，最后得到的 3 个候选元素就是答案了。

我们还可以将问题抽象成一个通用的情况：有 $k-1$ 个元素出现了超过 $\frac{1}{k}$ 次，求出这 $k-1$ 个元素。

```
// 时间复杂度 O(N), 空间复杂度 O(1)
int[] majorityNumber(ArrayList<Integer> nums, int k) {
    int[] candidates = new int[k];
    int[] count = new int[k];
    for (int i = 0; i < nums.size(); i++) {
        int cur = nums.get(i);
        for (j = 0; j < k; j++) {
            if (count[j] == 0) {
                candidates[j] = cur;
                count[j]++;
                break;
            } else if (candidates[j] == cur) {
                count[j]++;
                break;
            }
        }
        if (j >= k) {
            for(int idx=0;idx<k;idx++) {
                count[idx]--;
            }
        }
    }
    return candidates;
}
```

2.4 1 的数目

问题

从 1 到 N ，顺序写下这 N 个数字，比如 1,2,3,4,...，数一下这其中出现了多少次数字“1”。

1. 定义函数 $f(N)$ 返回 1 到 N 之间出现的 1 的个数，比如 $f(12) = 5$ 。
2. 满足条件 $f(N) = N$ 的最大的 N 是多少？

问题 1 解法

如果从 1 到 N 遍历每一个数字并且求出每个数字中包含了几个 1，这种做法的时间复杂度是 $O(N * \log_{10}(N))$ 。

我们可以这样考虑，我们首先观察个位数字是循环出现的，循环节是 10，每 10 个数

字出现一个 1。

$$\underbrace{0, \underline{1}, 2 \cdots, 8, 9}_{10}$$

再来观察一下十位数，这次每 100 个数字就会出现 10 个 1。

$$\underbrace{0, 1, 2 \cdots, 8, 9, \underline{10}, \underline{11}, \underline{12}, \underline{13}, \underline{14}, \underline{15}, \underline{16}, \underline{17}, \underline{18}, \underline{19}, 20, 21, \cdots, 98, 99}_{100}$$

同理，百位上每 1000 个数字出现 100 个 1...

下面，我们举个例子说明一下怎么求 1 出现的次数。N=1231，先看个位，总共会循环 123+1 次，每个循环有 1 个 1，再看十位，总共循环 12+1 次，每个循环有 10 个 1，再看百位，总共循环 1+1 次，每个循环有 100 个 1，最后是千位，总共循环 0.232 次，每个循环有 1000 个 1。

关键问题就是如何求循环次数，我们考虑第 K 位（个位是第 0 位），digit 为当前位上的数字，higher 为高位的数字，lower 为低位的数字。digit 和 1 的关系有三种：

1. digit=0，循环次数为 higher，出现次数为 $higher * 10^K$ 。
2. digit=1，循环次数为 higher，出现次数为 $higher * 10^K + lower + 1$ 。
3. digit>1，循环次数为 higher+1，出现次数为 $(higher + 1) * 10^K$ 。

```
// 时间复杂度 O(log10(N)), 空间复杂度 O(1)
int digitCounts(int n) {
    int ans = 0;
    int power = 1;
    int higher = 0, lower = 0;
    while (n != 0) {
        int digit = n % 10;
        higher = n / 10;
        ans += (higher + (digit > 1 ? 1 : 0)) * power;
        if (digit == 1) {
            ans += lower + 1;
        }
        lower += power * digit;
        power *= 10;
        n /= 10;
    }
    return ans;
}
```

问题 2 解法

无

扩展问题 1

现在给你你一个数字 $K, K \in [0, 1, 2, \dots, 8, 9]$ 。求 1 到 N 中出现 K 的次数。题目来源 Lint Code Digit Count。

解法

和原问题基本一致，只不过要特殊考虑一下 0，因为 0 不能作为数字开头。

扩展问题 2

对于其他进制表示方法，也可以试试，看看什么规律。例如二进制：

$$f(1) = 1$$

$$f(10) = 10 \quad (01, 10 \text{ 两个 } 1)$$

$$f(11) = 100 \quad (01, 10, 11 \text{ 四个 } 1)$$

解法

和十进制的做法基本一致，为了方便计算我们采用十进制来表示数字，比如上面第三个式子我们用十进制表示为 $f(3) = 4$ ，也就是输入输出都是十进制数，但是计算 1 的个数是采用二进制数。这里给出的代码采用的很多位运算。

```
// 时间复杂度  $O(\lg(N))$ ，空间复杂度  $O(1)$ 
int digitCountsBinary(int n) {
    int ans = 0;
    int wei = 0;
    int higher = 0, lower = 0;
    while (n != 0) {
        int digit = n & 0x01;
        higher = n >> 1;
        if (digit == 0) {
            ans += (higher << wei);
        } else {
            ans += ((higher << wei) | lower) + 1;
        }
        lower += (digit << wei);
        wei++;
        n >>= 1;
    }
    return ans;
}
```

2.5 寻找最大的 K 个数

问题

给你个很大的数组，求出这个数组最大的前 K 个数。

解法

拿到问题看到前 K 个，首先就想到了二叉堆，我们维护一个小根堆的二叉堆，用来保存最大的 K 个数。然后不断的遍历数组，每遇到一个新的元素就让它进堆，如果堆的大小大于 K 了，就把最小的元素删除，继续保持小根堆的性质不变。这样做既可以处理很大的数组，而且可以处理动态到来的元素，是一个在线算法。维护一个二叉堆的时间复杂度是 $O(\lg(K))$ ，总的复杂度是 $O(N * \lg(K))$ ，N 为数组长度。我们在具体的实现的时候，可以直接用 PriorityQueue 类来维护二叉堆。

```
// 时间复杂度  $O(N * \lg(K))$ ，空间复杂度  $O(K)$ 
int[] findMostKthNumber(int[] nums, int k) {
    PriorityQueue<Integer> queue = new PriorityQueue<Integer>(k + 1);
    for (int num : nums) {
        queue.offer(num);
        if (queue.size() > k) {
            queue.poll();
        }
    }
    int[] ans = new int[k];
    for (int i = 0; i < k; i++) {
        ans[i] = queue.poll();
    }
    return ans;
}
```

2.6 精确表达浮点数

问题

用分数表示小数。

$$0.9 = \frac{9}{10}$$

$$0.333(3) = \frac{1}{3}$$

$$0.1(2) = \frac{11}{90}$$

左边小数中括号部分为循环节。

解法

我们用 $0.\overbrace{a_1 \cdots a_k}^k \overbrace{(b_1 \cdots b_c)}^c$ 表示小数，其中非循环节部分长度为 k ，数值表示为 $\overline{a_1 \cdots a_k}$ ，循环节部分长度为 c ，数值表示为 $\overline{b_1 \cdots b_c}$ 。我们将最终要求的分数表示成 $\frac{A}{B}$

下面我们用公式推导出 A 和 B 。

$$\begin{aligned}
 x &= 0.\overbrace{a_1 \cdots a_k}^k \overbrace{(b_1 \cdots b_c)}^c \\
 10^k * x &= \overline{a_1 \cdots a_k}.\overline{(b_1 \cdots b_c)} \\
 10^{k+c} * x &= \overline{a_1 \cdots a_k b_1 \cdots b_c}.\overline{(b_1 \cdots b_c)} \\
 (10^{k+c} - 10^k) * x &= \overline{a_1 \cdots a_k b_1 \cdots b_c} - \overline{a_1 \cdots a_k} \\
 x = \frac{A}{B} &= \frac{\overline{a_1 \cdots a_k b_1 \cdots b_c} - \overline{a_1 \cdots a_k}}{10^k * (10^c - 1)} = \frac{\overline{a_1 \cdots a_k} * (10^c - 1) + \overline{b_1 \cdots b_c}}{10^k * (10^c - 1)} \\
 A &= \overline{a_1 \cdots a_k} * (10^c - 1) + \overline{b_1 \cdots b_c} \\
 B &= 10^k * (10^c - 1)
 \end{aligned}$$

```
// 时间复杂度 O(N)，空间复杂度 O(1)
int[] floatNubmer(String number) {
    int[] fraction = new int[2];
    int pre = 0, cycle = 0;
    int powerPre = 1, powerCycle = 1;
    boolean flag = false;
    for(int i=2; i<number.length() && number.charAt(i)!='.'; i++) {
        if(number.charAt(i)=='(') {
            flag = true;
            continue;
        }
        if(flag) {
            cycle = cycle * 10 + number.charAt(i) - '0';
            powerCycle *= 10;
        } else {
            pre = pre * 10 + number.charAt(i) - '0';
            powerPre *= 10;
        }
    }
    if(flag) {
        fraction[0] = pre * (powerCycle - 1) + cycle;
        fraction[1] = powerPre * (powerCycle - 1);
    } else {
        fraction[0] = pre;
        fraction[1] = powerPre;
    }
    int d = gcd(fraction[0], fraction[1]);
    fraction[0] /= d;
}
```

```
        fraction[1] /= d;
        return fraction;
    }

    int gcd(int a, int b) {
        if(a==0) return b;
        return gcd(b%a, a);
    }
```