会力和 学习 题库 竞赛 讨论 求职 商店 🗸

Plus 会员

讨论 / 技术交流 / 分享 | 从集合论到位运算, 常...

🕬 分享 | 从集合论到位运算,常见位运算技巧分类总结!

**灵茶山艾府 ♀** 发起于 2023-06-16 ● 最近编辑于 2024-11-04 ● 来自浙江

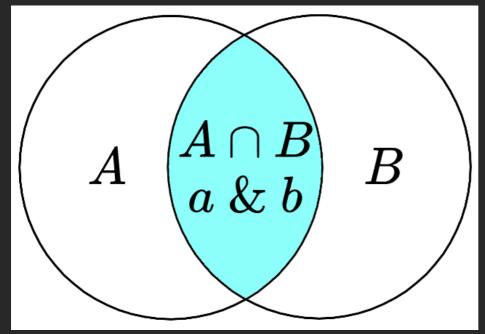


图:集合交、按位与之间存在某种联系。

# 前言

本文将扫清位运算的迷雾,在集合论与位运算之间建立一座桥梁。

在高中,我们学了集合论(set theory)的相关知识。例如,包含若干整数的集合 S = $\{0,2,3\}$ 。在编程中,通常用哈希表(hash table)表示集合。例如 Java 中的 HashSet , C++中的 std::unordered\_set 。

在集合论中,有交集 ∩、并集 ∪、包含于 ⊆ 等等概念。如果编程实现「求两个哈希表的交 集」,需要一个一个地遍历哈希表中的元素。那么,有没有效率更高的做法呢?

该二进制登场了。

集合可以用二进制表示,二进制**从低到高**第 i 位为 1 表示 i 在集合中,为 0 表示 i 不在集合 中。例如集合  $\{0,2,3\}$  可以用二进制数  $1101_{(2)}$  表示;反过来,二进制数  $1101_{(2)}$  就对应着 集合  $\{0,2,3\}$ 。

正式地说,包含非负整数的集合 S 可以用如下方式「压缩」成一个数字:

$$f(S) = \sum_{i \in S} 2^i$$

例如集合  $\{0,2,3\}$  可以压缩成  $2^0 + 2^2 + 2^3 = 13$ ,也就是二进制数  $1101_{(2)}$ 。

利用位运算「并行计算」的特点,我们可以高效地做一些和集合有关的运算。按照常见的应 用场景,可以分为以下四类:

- 1. 集合与集合
- 2. 集合与元素
- 3. 遍历集合
- 4. 枚举集合

### 一、集合与集合

其中 & 表示按位与, | 表示按位或, ⊕ 表示按位异或, ~ 表示按位取反。



#### LeetBook 推荐



LeetCode Cookbook



高效算法: 竞赛、应...



零起步学算法

- △ 相关讨论
- ♠ 2024 Rewind | 开启你的年...
- 一鼓作气 | 2024 扣得精彩...
- ぐ [力扣刷题攻略] Re: 从零...
- 「力扣编辑器」使用说明
- 🥟 求助 | 本人双非研二 本科...



**口流间** 力扣(LeetCode)App 随时随地参与讨论

App 扫一扫 即刻参与

两个集合的「对称差」是只属于其中一个集合,而不属于另一个集合的元素组成的集合,也就是不在交集中的元素组成的集合。

术语	集合	位运算	集合示例	位运算示例
交集	$A\cap B$	a&b	$\{0,2,3\} \ \cap \{0,1,2\} \ = \{0,2\}$	1101 & 0111 = 0101
并集	$A \cup B$	$a\mid b$	$\{0,2,3\} \ \cup \{0,1,2\} \ = \{0,1,2,3\}$	1101   0111 = 1111
对称差	$A \Delta B$	$a\oplus b$	$\{0,2,3\}$ $\Delta \ \{0,1,2\}$ $= \{1,3\}$	$1101$ $\oplus 0111$ $= 1010$
差	$A\setminus B$	$a\&\sim b$	$\{0,2,3\} \ ackslash \{1,2\} \ = \{0,3\}$	1101 & 1001 = 1001
差 (子集)	$A\setminus B,\ B\subseteq A$	$a\oplus b$	$\{0,2,3\} \ ackslash \{0,2\} \ = \{3\}$	$1101$ $\oplus 0101$ $= 1000$
包含于	$A\subseteq B$	$a\&b=a \ a\mid b=b$	$\{0,2\}\subseteq \{0,2,3\}$	$0101\&1101 = 0101$ $0101 \mid 1101 = 1101$

注 1: 按位取反的例子中,仅列出最低 4 个比特位取反后的结果,即 0110 取反后是 1001。

注 2: 包含于 (判断子集) 的两种位运算写法是等价的,在编程时只需判断其中任意一种,此外,还可以用  $(a \& \hat{b}) == 0$  判断,如果成立,也表示  $A \in B$  的子集,

注 3:编程时,请注意运算符的优先级。例如 == 在某些语言中优先级比位运算更

# 二、集合与元素

通常会用到移位运算。

其中 << 表示左移, >> 表示右移。

注: 左移 i 位相当于乘以  $2^i$ ,右移 i 位相当于除以  $2^i$ 。

术语	集合	位运算	集合示例	位运算示例
空集	Ø	0		
单元素集合	$\{i\}$	1 << i	{2}	1 << 2
全集	$U= \{0,1,2,\ldots,n-1\}$	(1 << n) - 1	$\{0,1,2,3\}$	(1 << 4) - 1
补集	$\complement_U S = U \setminus S$	$((1 \mathrel{<<} n) - \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ $	$egin{aligned} U = \ \{0,1,2,3\} \ \mathbb{C}_U\{1,2\} = \ \{0,3\} \end{aligned}$	1111 ⊕ 0110 = 1001
属于	$i \in S$	$(s \gt\gt i) \ \& \ 1 = 1$	$2\in\{0,2,3\}$	(1101 >> 2) & 1 =

术语	集合	位运算	集合示例	位运算示例
不属于	i otin S	$(s \gt\gt i) \ \& \ 1 = 0$	$1\notin\{0,2,3\}$	(1101 >> 1) & 1 = 0
添加元素	$S \cup \{i\}$	$s \mid (1 \mathrel{<\!\!\!<} i)$	$\{0,3\}\cup\{2\}$	1001   (1 << 2)
删除元素	$S\setminus\{i\}$	$s\&\sim (1  ext{ << } i)$	$\{0,2,3\}\setminus\{2\}$	$1101\&\sim \ (1<<2)$
删除元素(一定在集合中)	$S\setminus\{i\},\ i\in S$	$s \oplus (1 \mathrel{<\!\!\!<} i)$	$\{0,2,3\}\setminus\{2\}$	1101 ⊕ (1 << 2)
删除最小元素		s&(s-1)		见下

特别地,如果 s 是 2 的幂,那么 s&(s-1)=0。

此外,编程语言提供了一些和二进制有关的库函数,例如:

- 计算二进制中的 1 的个数,也就是**集合大小**;
- 计算二进制长度,减一后得到集合最大元素;
- 计算二进制尾零个数,也就是集合最小元素。

调用这些函数的时间复杂度都是  $\mathcal{O}(1)$ 。

术语	Python	Java	C++	Go
集合大小	s.bit_count ()	Integer.bit Count(s)	builtin_p opcount(s)	bits.OnesCo
二进制长度	s.bit_lengt h()	32- Integer.num berOfLeadin gZeros(s)	lg(s)+1	bits.Len(s)
集合最大元素	s.bit_lengt h()-1	31- Integer.num ber0fLeadin gZeros(s)	lg(s)	bits.Len(s)
集合最小元素	(s&- s).bit_leng th()-1	Integer.num berOfTraili ngZeros(s)	builtin_c tz(s)	bits.Traili ngZeros(s)

请特别注意 s=0 的情况。对于 C++ 来说, $_{-}$ 1g(0) 和  $_{-}$ builtin\_ctz(0) 是未定义行 为。其他语言请查阅 API 文档。

此外,对于 C++ 的 long long ,需使用相应的 \_\_builtin\_popcountll 等函数,即函数 名后缀添加 11 (两个小写字母 L) 。  $_{-1g}$  支持 long long 。

特别地,只包含最小元素的子集,即二进制最低1及其后面的0,也叫lowbit,可以用 s & -s 算出。举例说明:

(`s)+1 = 010100 // 根据补码的定义,这就是 -s => s 的最低 1 左侧取反,右侧不变 s & -s = 000100 // lowbit

# 三、遍历集合

设元素范围从 0 到 n-1,枚举范围中的元素 i,判断 i 是否在集合 s 中。

也可以直接遍历集合 s 中的元素:不断地计算集合最小元素、去掉最小元素,直到集合为空。

```
Python3 | Java | C++ | Go

for (int t = s; t; t &= t - 1) {
    int i = __builtin_ctz(t);
    // 处理 i 的逻辑
}
```

# 四、枚举集合

#### §4.1 枚举所有集合

设元素范围从 0 到 n-1,从空集  $\emptyset$  枚举到全集 U:

```
Python3 | Java | C++ | Go

for (int s = 0; s < (1 << n); s++) {
    // 处理 s 的逻辑
}
```

#### §4.2 枚举非空子集

设集合为 s, **从大到小**枚举 s 的所有**非空**子集 sub:

为什么要写成 sub = (sub - 1) & s 呢?

暴力做法是从 s 出发,不断减一,直到 0。但这样做,中途会遇到很多并不是 s 的子集的情况。例如 s=10101 时,减一得到 10100,这是 s 的子集。但再减一就得到 10011 了,这并不是 s 的子集,下一个子集应该是 10001。

把所有的合法子集按顺序列出来,会发现我们做的相当于「压缩版」的二进制减法,例如

```
10101 \rightarrow 10100 \rightarrow 10001 \rightarrow 10000 \rightarrow 00101 \rightarrow \cdots
```

如果忽略掉 10101 中的两个 0, 数字的变化和二进制减法是一样的, 即

```
111 
ightarrow 110 
ightarrow 101 
ightarrow 100 
ightarrow 011 
ightarrow \cdots
```

如何快速跳到下一个子集呢?比如,怎么从10100跳到10001?

- 普通的二进制减法,是 10100 1 = 10011,也就是把最低位的 1 变成 0,同时把最低位的 1 右边的 0 都变成 1。
- 压缩版的二进制减法也是类似的,对于  $10100 \rightarrow 10001$ ,也会把最低位的 1 变成 0,对于最低位的 1 右边的 0,并不是都变成 1,只有在 s=10101 中的 1 才会变成 1。怎么做到?减一后 & 10101 就行,也就是 (10100-1) & 10101=10001。

### §4.3 枚举子集 (包含空集)

如果要从大到小枚举 s 的所有子集 sub (从 s 枚举到空集  $\varnothing$ ) ,可以这样写:

```
Python3 | Java | C++ | Go
```

原理是当 sub=0 时(空集),再减一就得到 -1,对应的二进制为  $111\cdots 1$ ,再 &s 就得到了 s。所以当循环到 sub=s 时,说明最后一次循环的 sub=0(空集),s 的所有子集都枚举到了,退出循环。

i: 还可以枚举全集 U 的所有大小**恰好**为 k 的子集,这一技巧叫做 Gosper's Hack,具体请看 视频讲解。

### §4.4 枚举超集

如果  $T \in S$  的子集,那么称  $S \in T$  的**超集** (superset) 。

枚举超集的原理和上文枚举子集是类似的,这里通过**或运算**保证枚举的集合 S 一定包含集合 T 中的所有元素。

枚举 S,满足 S 是 T 的超集,也是全集  $U = \{0,1,2,\ldots,n-1\}$  的子集。

```
Python3 | Java | C++ | Go

for (int s = t; s < (1 << n); s = (s + 1) | t) {
    // 处理 s 的逻辑
}
```

## 关联题单

- 位运算题单
- 数据结构题单中的「前缀异或和」
- 动态规划题单中的「状压 DP」

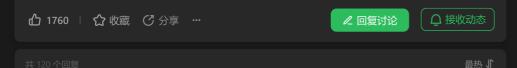
### 分类题单

如何科学刷题?

- 1. 滑动窗口与双指针 (定长/不定长/单序列/双序列/三指针)
- 2. 二分算法 (二分答案/最小化最大值/最大化最小值/第K小)
- 3. 单调栈(基础/矩形面积/贡献法/最小字典序)
- 4. 网格图 (DFS/BFS/综合应用)
- 5. 位运算(基础/性质/拆位/试填/恒等式/思维)
- 6. 图论算法 (DFS/BFS/拓扑排序/最短路/最小生成树/二分图/基环树/欧拉路径)
- 7. 动态规划 (入门/背包/状态机/划分/区间/状压/数位/数据结构优化/树形/博弈/概率期望)
- 8. 常用数据结构 (前缀和/差分/栈/队列/堆/字典树/并查集/树状数组/线段树)
- 9. 数学算法 (数论/组合/概率期望/博弈/计算几何/随机算法)
- 10. 贪心与思维 (基本贪心策略/反悔/区间/字典序/数学/思维/脑筋急转弯/构造)
- 11. 链表、二叉树与一般树(前后指针/快慢指针/DFS/BFS/直径/LCA)
- 12. 字符串(KMP/Z函数/Manacher/字符串哈希/AC自动机/后缀数组/子序列自动机)

我的题解精选 (已分类)

欢迎关注 B站@灵茶山艾府



◎ 天赐细莲 ⑩

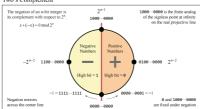
来自江苏。 2023-06-16

分享一张图

# **Binary Fundamentals**

#### Powers of Two

Power	Decimal	Hexadecimal
21	2	0x00000002
2 <sup>2</sup>	4	0x00000004
23	8	0x00000008
24	16	0x00000010
25	32	0x00000020
26	64	0x00000040
27	128	0x00000080
28	256	0x00000100
29	512	0x00000200
210	1024	0x00000400
211	2048	0x00000800
212	4096	0x00001000



NOT			
NOI	×	~×X	
	9	1	
Bitwise NOT	1	0	
~x			

Unary	Binary
~x = -x - 1	~(x & y) = ~x   ~y
-x = ~x + 1	~(x   y) = ~x & ~y
-~x = x + 1	~(x ^ y) = { ~x ^ y x ^ ~y
~-x = x - 1	(x ^ ~y) - ( x ^ ~y

#### Floating-Point



-0.0

Special Floating-Point Value	Half	Float	Double
+0.0	9x9999	0x00000000	0×00000000_00000000
+1.0	0x3C00	0x3F800000	0x3FF00000_00000000
Positive infinity	0x7C00	0x7F800000	0x7FF00000_00000000
Smallest positive normalized value	0x0400	0×00800000	0×00100000_00000000
Upper limit of non-integer values	9x6499	0x4B000000	0x43300000_000000000
Largest representable positive value	0x7BFF	0x7F7FFFFF	0x7FEFFFFF FFFFFFF

#### Binary Logical Operations A N D x y x & y

(2) [3] [1]					(112				
AND	0	0	0		UK	0	0	0	ı
Ritwise AND	0	1	0		Bitwise OR	0	1	1	ı
	1	0	0			1	0	1	ı
x & y	1	1	1		x   y	1	1	1	ı
				_					4
NAND	X Ø	y 0	~(x & y)		NOR	× 0	y 0	~(x   y)	
	0	1	1			0	1	9	ı
Not AND	1	0	1		Not OR	1	0	9	ı
~(x & y)	1	1	0		~(x   y)	1	1	9	ı
				_					
ANDC	×	у	x & ~y		ORC	Х	у	x   ~y	ı
ANDC	0	0	0		UKC	0	0	1	ı
AND with complement	0	1	0		OR with complement	0	1	0	ı
	1	0	1			1	0	1	ı
x & ~y	1	1	0		x   ~y	1	1	1	ı

**XNOR** 

#### Mask Creation

Formula	Operation / Effect	Illustration
~x   (x - 1)	Create mask for all bits other than lowest 1 bit. 900 900 becomes 111 111.	1 0 1 1 1 0 0 0
x   ~(x + 1)	Create mask for all bits other than lowest 0 bit. 111121 is unchanged.	0 1 1 0 0 1 1 1
x   -x	Create mask for bits left of lowest 1 bit, inclusive. 000000 is unchanged.	1 0 1 1 1 0 0 0
x ^ -x	Create mask for bits left of lowest 1 bit, exclusive. 000000 is unchanged.	1 0 1 1 1 0 0 0
~x   (x + 1)	Create mask for bits left of lowest 0 bit, inclusive.  111111 becomes 800900.	0 1 1 0 0 1 1 1
~x ^ (x + 1)	Create mask for bits left of lowest 0 bit, exclusive.  111111 becomes 800000.	0 1 1 0 0 1 1 1
x ^ (x - 1)	Create mask for bits right of lowest 1 bit, inclusive. 000000 becomes 111111.	0 0 0 0 1 1 1 1
~x & (x - 1)	Create mask for bits right of lowest 1 bit, exclusive.	0 0 0 0 0 1 1 1
x ^ (x + 1)	Create mask for bits right of lowest 0 bit, inclusive.  111111 is unchanged.	0 1 1 0 0 1 1 1
x & (~x - 1)	Create mask for bits right of lowest 0 bit, exclusive.  111111 is unchanged.	0 1 1 0 0 1 1 1

**XOR** 

Formula	Operation / Effect	Illustration
x & (x - 1)	Clear lowest 1 bit.  If result is zero, then x is zero or 2 <sup>k</sup> .  990990 is unchanged.	1 0 1 1 0 0
x   (x + 1)	Set lowest 0 bit. 111111 is unchanged.	0 1 1 0 0 1 1
x   (x - 1)	Set all bits to right of lowest 1 bit. 000-000 becomes 111111.	1 0 1 1 1 0 0
x & (x + 1)	Clear all bits to right of lowest 0 bit.  If result is zero, then x is zero or 2 <sup>k</sup> = 1.  111···111 becomes 900···900.	0110011
x & -x	Extract lowest 1 bit. 990999 is unchanged.	0 0 0 0 1 0 0
~x & (x + 1)	Extract lowest 0 bit (as a 1 bit). 111111 becomes 000000.	0 1 1 0 0 1 1









**心**添加回复



○ 灵茶山艾府 ※ · 楼主





② 分享一张图

谢谢分享。位运算的技巧多到可以写一本书,本文列出的是做题时会经常用到的位运算技 巧。

△ 134 | ○ 回复 ☆ 收藏 ♂ 分享 …

**/** 添加回复



灵神太强了

△ 32 │ ○ 回复 1 ☆ 收藏 ② 分享 …

**/** 添加回复



「压缩版」的二进制减法太帅了! 只是 但对于这个 1 右边的 0,只保留在 10101 中的 1 这句话太绕了。

我啰嗦一下,意思是,把10100的最低位1变0,它的后面有两位00,都是0。这时候按照 普通二进制, 会把这两个 00 都变成 11, 如果按照压缩版, 就只把原来集合里有的 1 变成 1 (因为求的是子集), 其余的还是0,原有的集合是10101,最后两位是01,所以只保留 01。综合起来就是 10100 先变 10000, 然后保留 01, 变成 10001。

这样做的效果是从 10100 直接跳到 10001, 把中间的 10011 和10010 忽略掉了(普通减法 顺序是 10100 - 10011 - 10010 - 10001) , 因为10011 和 10010不是有效的子集。

**△** 20 │ ○ 回复 2 **◇** 收藏 **○** 分享 ····

**/** 添加回复

🦒 Automation垃圾专业

应用数学中混进了一个理论数学

△8 □ ○回复 ☆ 收藏 ♂ 分享 …

**心**添加回复

**愛** 灵剑2012

枚举固定大小的子集能用bit\_count相关指令的话好像还挺容易的,是不是就是每次加lowbit 然后在最低位补全1的数量就行

△8 □ ○回复 ◇ 收藏 ② 分享 …

**儿**添加回复

🦀 悟🛊 🤍

我的神!!!!!

△8 | ○回复 ☆ 收藏 ② 分享 …

**/** 添加回复

🎇 KZHU 😵

② 分享一张图

感谢大佬分享 https://terathon.com/binary\_fund.pdf 附上链接大家直接下载下来当壁纸 吧曾

△ 6 □ ○ 回复 ☆ 收藏 ② 分享 …

**心**添加回复

○ ? · 对茶山艾府 ※ · 楼主

(1) 请教灵神,怎么理解位运算的「并行计算」这个特点。

二进制每一位是互相独立的。比如 AND 运算,在某个比特位上,两个数必须都是 1,结果 才能是 1, 把这个运算规则应用到每个比特位就能得到结果。例如 int 有 32 个比特, 如果一 个个地去算就要算32次,但是位运算只需要一次就能全部算完。

**△** 5 │ **○** 回复 **◇** 收藏 **○** 分享 ····

**/** 添加回复

yuukilee

真正的英雄!

遍历子集泰裤辣



沪ICP备18019787号-20 <u>9</u> 沪公网安备31010702007420号 沪ICP证B2-20180578 LeetCode力扣·证照中心 ト海市互联网违法和不良信息举报中小》 中国互联网违法和不良信息举报中小》