基数排序

- 不通过待排序数据元素之间的比较
- 根据关键字本身的性质进行排序 分配排序,桶排序,基数排序
- 例1: 名字按照字典序排序
- 例2: 1—m之间的整数排序

基数排序

- 例3:一年的全国高考考生人数为500万,分数100--900, 没有小数,把这500万考生按照分数排序。
- > 可能的分数为801个,创建801个"桶",每个桶对应 一个分值
- » 依次查看每个学生的成绩,将其放入对应其分值的<mark>桶</mark>
- > 从900分对应的那个桶开始依次收集,即完成排序
- 如果待排序的数据元素的取值是从0到2亿?
- » 那就要分配2亿个桶了,这"太狼费"的,所以桶排 序有其局限性,适合元素 关键字值集合并不大的情况

桶排序

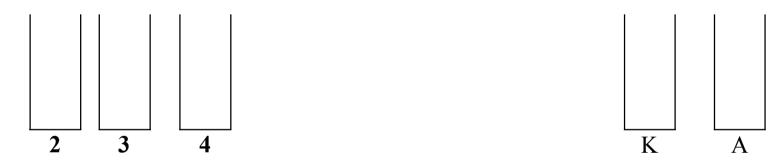
- 桶排序的基本操作:
- > 分配关键字:根据每个数据元素关键字的值将其分配到相应的桶中
- > 每个"桶"内排序数据元素均匀分布,则每个桶中数据元素个数均匀
- > 收集桶
- 不能象比较排序那样以统一的数据元素之间的"比较"次数衡量"工作量"
- 分析每一步骤地工作进行时间复杂度的分析

桶排序要求: 关键字的取值范围或结构已知

基数排序

- 一张牌有两个关键字组成:花色(^<♥<♣<◆)+面值(2<3<4<...<A)。假如一张牌的大小首先被花色决定,同花色的牌由数字决定。
- 52张牌排序有两种算法:
 - (1)首先按照面值对所有牌进行排序,然后按照花色再次对所有牌进行排序。
 - (2)首先按照花色将所有牌分成4组。然后同组的牌(同花色)再按照面值进行排序。

(1)首先按照面值对所有牌进行排序,然后按照花色再次对所有牌进行排序。 牌的面值13个,准备13个空桶,分别对应13个面值



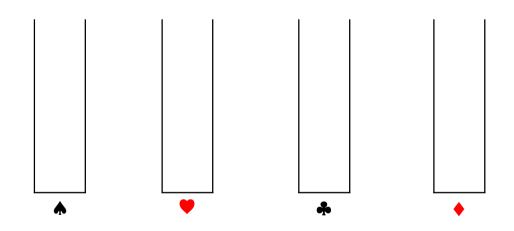
依次查看52张牌,根据其面值将其放入对应的桶中

(1)首先按照面值对所有牌进行排序,然后按照花色再次对所有牌进行排序。 牌的面值13个,准备13个空桶,分别对应13个面值



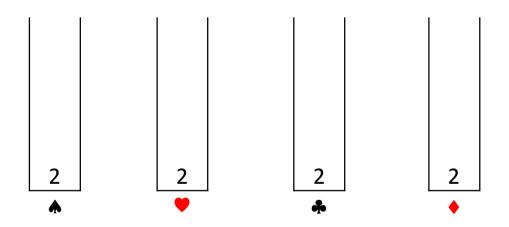
依次查看52张牌,根据其面值将其放入对应的桶中----第一次分配 从"2"号开始依次收集,将每个桶中的牌顺次倒出,52张牌现在的顺序是4 张2,4张3,4张4,……,4张K,4张A----第一次收集

根据第一次收集的结果然后按照花色再次对所有牌进行排序。花色有4个, 准备4个空桶,分别对应4个花色



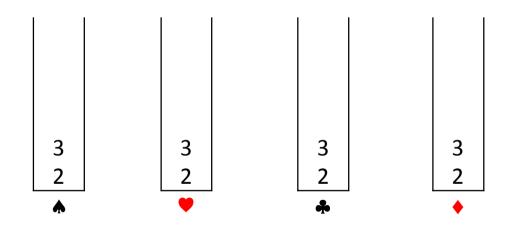
依次查看第一次收集得到52张牌,根据花色将其放入对应的桶中----第二次分配

根据第一次收集的结果然后按照花色再次对所有牌进行排序。花色有4个,准备4个空桶,分别对应4个花色



依次查看第一次收集得到52张牌,根据花色将其放入对应的桶中----第二次分配

根据第一次收集的结果然后按照花色再次对所有牌进行排序。花色有4个,准备4个空桶,分别对应4个花色



依次查看第一次收集得到52张牌,根据花色将其放入对应的桶中----第二次分配

收集桶的方式: 最先进入一个桶的"2"最先倒出----实现 时考虑用队列模拟桶

基数排序-方法1

根据第一次收集的结果然后按照花色再次对所有牌进行排序。花色有4个,准备4个空桶,分别对应4个花色

扑克牌的排序是 根据其2个值决定 的:花色和面值, 且花色的优先级 高于面值

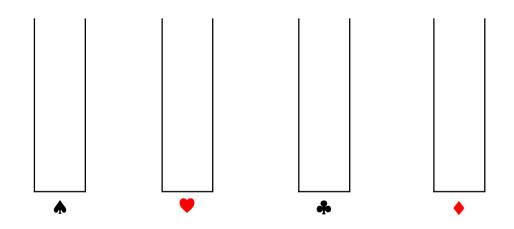
A K	
•	
4	
3	
2	
	•

	A K	
	•	
	4	
	3	
	2	
,	•	

A K	
4	
3	
2	
•	

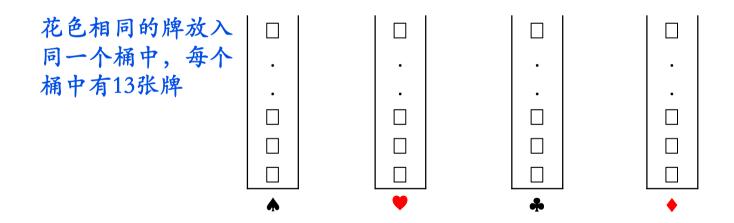
依次查看第一次收集得到52张牌,根据花色将其放入对应的桶中----第二次分配从"♠"号桶开始依次收集,将每个桶桶口封死,桶底打开,顺次倒出每一个桶中的牌,52张牌就排好序了: 13张♠从2到A, 13张♥从2到A, 13张♣从2到A, 13张◆从2到A, 13张◆从2到A, 13张◆从2到A, 13张◆从2到A, 13张◆

(2)首先按照花色将所有牌分成4组。然后同组的牌(同花色)再按照面值进行排序。花色有4个,准备4个空桶,分别对应4个花色



依次查看52张牌,根据花色将其放入对应的桶中----第一次分配

(2)首先按照花色将所有牌分成4组。然后同组的牌(同花色)再按照面值进行排序。花色有4个,准备4个空桶,分别对应4个花色



依次查看52张牌,根据花色将其放入对应的桶中----第一次分配

第一次收集? 能否和前面的方法一样将所有相依没收集在一起?

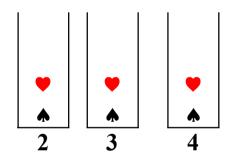
岩4个花色的桶依次一起收集,则第一次收集的52张牌为:13张♠,13张♥,13 张♣,13张♦。牌的面值13个,准备13个空桶,分别对应13个面值

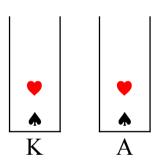


岩4个花色的桶依次一起收集,则第一次收集的52张牌为:13张♠,13张♥,13 张♣,13张♦。牌的面值13个,准备13个空桶,分别对应13个面值

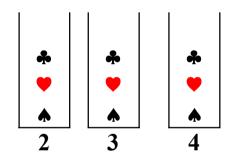


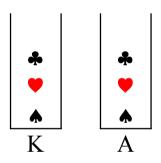
老4个花色的桶像次一起收集,则第一次收集的52张牌为:13张★,13张♥,13 张★,13张♦。牌的面值13个,准备13个空桶,分别对应13个面值



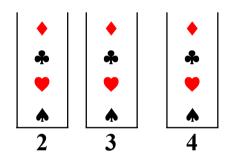


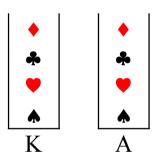
老4个花色的桶像次一起收集,则第一次收集的52张牌为:13张★,13张♥,13 张★,13张♦。牌的面值13个,准备13个空桶,分别对应13个面值





老4个卷色的桶像次一起收集,则第一次收集的52张牌为:13张★,13张♥,13 张★,13张♦。牌的面值13个,准备13个空桶,分别对应13个面值



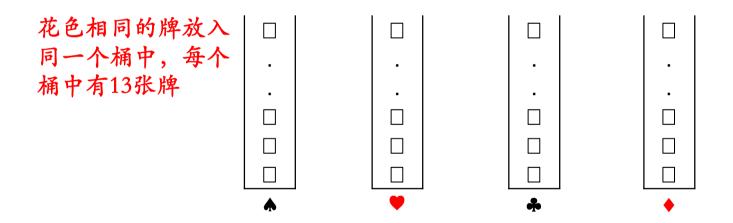


依次查看第一次收集得到的52张牌,根据其面值将其放入对应的桶中

第二次收集? 一起收集得不到正确的排序结果。

基数排序----正确的方法一

(2)首先按照花色将所有牌分成4组。然后同组的牌(同花色)再按照面值进行排序。花色有4个,准备4个空桶,分别对应4个花色



依次查看52张牌,根据花色将其放入对应的桶中----第一次分配

第一次收集? 不一起收集。先收集《桶,对收集的结果按照面值进行桶排序,再收集》桶,对收集的结果按随面值进行桶排序,……

基数排序

- 基数排序是一种借助"多关键字排序"的思想来实现 "单关键字排序"的内部排序算法。
- n 个记录的序列 $\{R_1, R_2, \dots, R_n\}$,对关键字 $(K_i^0, K_i^1, \dots, K_i^{d-1})$ 有序是指:对于序列中任意两个记录 R_i 和 R_j ($1 \le i < j \le n$) 都满足下列(词典)有序关系:

 $(K_i^0, K_i^1, \dots, K_i^{d-1}) < (K_j^0, K_j^1, \dots, K_j^{d-1})$

K⁰ 被称为"最主"位关键字

Kd-1 被称为"最次"位关键字



- □ 2种基数排序方法:
- 1. 最高位优先MSD法
- 2. 最低位优先LSD法: 先对Kd-1进行桶排序排序,再对 Kd-2 进行桶排序, ...··, 依次类推, 直至最后对最高位关键字桶排序完成为止。
- 通常采用低位优先
- 以下以低位优先为例介绍基数排序

基数排序

- 例 {369,367,167,239,237,138,230,139}采用低位优先基数排序,首先拆成3个关键字:个位、十位、百位,然后从个位开始进行以下3趟基数排序:
- 1. 首先按其"个位数"取值分配到"0,1,…,9"10个桶,之后按从0至9的顺序将它们"收集"在一起;
- 2. 然后按其"十位数"取值分配到"0,1,…,9" 10个桶,之后再按 从0至9的顺序将它们"收集"在一起;
- 3. 最后按其"百位数"重复一遍上述操作。

基数排序的实现

- 采用链表作存储结构--链式基数排序,用队列模拟"桶", 具体作法为:
- 1. 待排序记录以指针相链,构成一个链表;
- 2. "分配"时,按当前"关键字位"所取值,将记录分配到不同的"链队列"(桶)中,每个队列(桶)中记录的"关键字位"相同;
- 3. "收集"时,按当前关键字位取值从小到大将各队列(桶)首尾相链成一个链表;
- 4. 对每个关键字位均重复2和3两步。

 $p \rightarrow 369 \rightarrow 367 \rightarrow 167 \rightarrow 239 \rightarrow 237 \rightarrow 138 \rightarrow 230 \rightarrow 139$ 进行第一次分配 第一次分配+第一次收集= 第一趟基数排序, 结果待排 $f[0] \rightarrow 230 \leftarrow r[0]$ 序数据按个位数升序排列 $f[7] \longrightarrow 367 \longrightarrow 167 \longrightarrow 237 \longleftarrow r[7]$ $f[8] \rightarrow 138 \leftarrow r[8]$ $f[9] \rightarrow 369 \rightarrow 239 \rightarrow 139 \leftarrow r[9]$

进行第一次收集

$$p \rightarrow 230 \rightarrow 367 \rightarrow 167 \rightarrow 237 \rightarrow 138 \rightarrow 369 \rightarrow 239 \rightarrow 139$$

 $p \rightarrow 230 \rightarrow 367 \rightarrow 167 \rightarrow 237 \rightarrow 138 \rightarrow 369 \rightarrow 239 \rightarrow 139$

进行第二次分配

$$f[3] \rightarrow 230 \rightarrow 237 \rightarrow 138 \rightarrow 239 \rightarrow 139 \leftarrow r[3]$$

$$f[6] \rightarrow 367 \rightarrow 167 \rightarrow 369 \leftarrow r[6]$$

进行第二次收集

p→230→237→138→239→139 →367→167→369 第二次分配+第二次收集=第二趟基数排序,结果待 排序数据按个和十位数升序排列

$$p \longrightarrow 230 \longrightarrow 237 \longrightarrow 138 \longrightarrow 239 \longrightarrow 139 \longrightarrow 367 \longrightarrow 167 \longrightarrow 369$$

进行第三次分配

$$f[1] \rightarrow 138 \rightarrow 139 \rightarrow 167 \leftarrow r[1]$$

$$f[2] \rightarrow 230 \rightarrow 237 \rightarrow 239 \leftarrow r[2]$$

$$f[3] \rightarrow 367 \rightarrow 369 \leftarrow r[3]$$

进行第三次收集之后便得到记录的有序序列

p→138→139→167→230→237→239→367→369 第三次分配+第三次收集=第二趟基数排序,结果待 排序数据按个、十位和百位数升序排列



基数排序

- "分配"和"收集"的实际操作仅为修改链表中的指 针和设置队列的头、尾指针;
- 为查找使用,该链表尚需应用算法Arrange 将它调整 为有序表。
- 时间复杂度为O(d(n+rd))。其中:
- › 分配为O(n), 收集为O(rd)(rd为"基"), d为"分配-收集"的 趟数
- 空间复杂度为O(rd)

数组下标

0	1	2	3	4	5	6	7
369	367	167	239	237	138	230	139

指针: 直接后继的地址=直接后继所

	指针: 直接后继的地址=直接后继所 在数组元素的下标 1 2 3 4 5 6 7															1下标		
0 1 2						3 4 5						6		7				
3	69	1	36	7	2	167	3	239	4	237	5	138	6	230	7	139		



							在	致 组刀	山系	的「初	<u> </u>				4		
101		0	1		2		3		4		5		6		7		
r		369	1	367	2	167	3	239	4	237	5	138	6	230	7	139	-1

▲ 数组下标

-1

p -1

p -1

p -1

				p) \										
0		1		2		3		4		5		6		7	
369	1	367	2	167	3	239	4	237	5	138	6	230	7	139	-1

				p) \										
0		1		2		3		4		5		6		7	
369	1	367	2	167	3	239	4	237	5	138	6	230	7	139	-1

						<u>p</u>	1							_	
0		1		2		3		4		5		6		7	
369 1		367	2	167	3	239	4	237	5	138	6	230	7	139	-1

		_				p	1								
0		1		2		3		4		5		6		7	
369	3	367	2	167	3	239	4	237	5	138	6	230	7	139	-1

_								_p \	1	-					
0 1 2 3						4		5		6		7			
369	3	367	2	167	3	239	4	237	5	138	6	230	7	139	-1

_								_p \	1	-					
0 1 2 3						4		5		6		7			
369	3	367	2	167	4	239	4	237	5	138	6	230	7	139	-1

								_p \	1						
0 1 2 3						4		5		6		7			
369	3	367	2	167	4	239	4	237	5	138	6	230	7	139	-1

										<u>b/</u>					
0		1		2		3		4		5		6		7	
369	3	367	2	167	4	239	4	237	5	138	6	230	7	139	-1

										<u>b/</u>					
0		1		2		3		4		5		6		7	
369	3	367	2	167	4	239	4	237	5	138	6	230	7	139	-1

								_				<u> </u>	1		_
0		1		2		3		4		5		6		7	
369	3	367	2	167	4	239	4	237	5	138	6	230	7	139	-1

								_				<u> </u>	1		_
0		1		2		3		4		5		6		7	
369	3	367	2	167	4	239	4	237	5	138	6	230	7	139	-1

								_				_		h /	_
0		1		2		3		4		5		6		7	
369	3	367	2	167	4	239	4	237	5	138	6	230	7	139	-1

														ΡŢ	_
0		1		2		3		4		5		6		7	
369	3	367	2	167	4	239	4	237	5	138	6	230	7	139	-1

														ΡŢ	_
0		1		2		3		4		5		6		7	
369	3	367	2	167	4	239	7	237	5	138	6	230	7	139	-1

														ΡŢ	_
0		1		2		3		4		5		6		7	
369	3	367	2	167	4	239	7	237	5	138	6	230	7	139	-1

														ΡŢ	_
0 1 2 3 4							5		6		7				
369	3	367	2	167	4	239	7	237	5	138	6	230	7	139	-1

														b /	
0 1 2 3 4 5								6		7					
369	3	367	2	167	4	239	7	237	5	138	6	230	7	139	-1

$$f[\mathbf{0}] \rightarrow 23\mathbf{0} \leftarrow r[\mathbf{0}]$$

										_		-		ΡŢ	_
0 1 2 3 4								5		6		7			
369	3	367 2 167 4 239 7		7	237	5	138	6	230	7	139	-1			

								_		-		-		ΡŢ	_
0 1 2 3 4							5		6		7				
369	3	3 367 2 167 4 239 7		7	237	5	138	6	230	7	139	-1			

														ΡŢ	_
0 1 2 3 4							5		6		7				
369	3	367	2	167	4	239	7	237	5	138	6	230	7	139	-1

														$p \setminus$	
0 1 2 3 4 5								6		7					
369	3	367	67 2 167 4 239				7	237	5	138	6	230	7	139	-1
				1				1							

$$f[7] \rightarrow 367 \rightarrow 167 \rightarrow 237 \leftarrow r[7]$$

														р\	
0 1 2 3 4						5		6		7					
369	3	367	2	167	4	239	7	237	5	138	6	230	7	139	-1

$$f[8] \rightarrow 138 \leftarrow r[8]$$

															$p \setminus$	
	0		1		2		3	3	4		5		6		7	
I	369	3	367	2	167	4	239	7	237	5	138	6	230	7	139	-1
ĺ																

$$f[9] \rightarrow 369 \rightarrow 239 \rightarrow 139 \leftarrow r[9]$$

												h /	1		
0 1 2 3 4 5 6								7							
369	3	367	2	167	167 4 239 7		7	237	5	138	6	230	7	139	-1

												h /	1		
0 1 2 3 4 5 6								7							
369	3	367	2	167	167 4 239 7		7	237	5	138	6	230	7	139	-1

												h /	1		_
0 1 2 3 4 5 6										7					
369	3	367 2 167 4 239 7		7	237	5	138	6	230	7	139	-1			

								_		_		b 7			
0 1 2 3 4 5 6										7					
369	3	367	57 2 167 4 239 7		7	237	5	138	6	230	1	139	-1		

$$\mathbf{p} \rightarrow 230 \rightarrow 367 \rightarrow 167 \rightarrow 237$$

p \														_	
0		1		2		3		4		5		6		7	
369	3	367	2	167	4	239	7	237	5	138	6	230	1	139	-1

$$\mathbf{p} \rightarrow 230 \rightarrow 367 \rightarrow 167 \rightarrow 237$$

P \														_	
0		1		2		3		4		5		6		7	
369	3	367	2	167	4	239	7	237	5	138	6	230	1	139	-1

$$\mathbf{p} \rightarrow 230 \rightarrow 367 \rightarrow 167 \rightarrow 237 \rightarrow 138$$

P \															
0		1		2	2		3		4		5		6		
369	3	367	2	167	4	239	7	237	5	138	0	230	1	139	-1

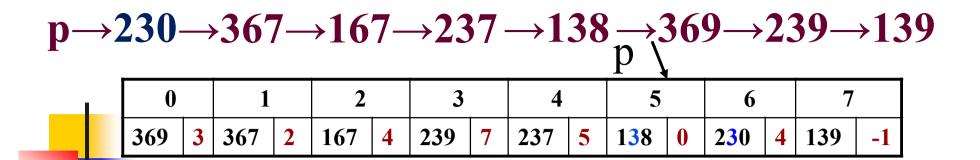
$$p \rightarrow 230 \rightarrow 367 \rightarrow 167 \rightarrow 237 \rightarrow 138$$

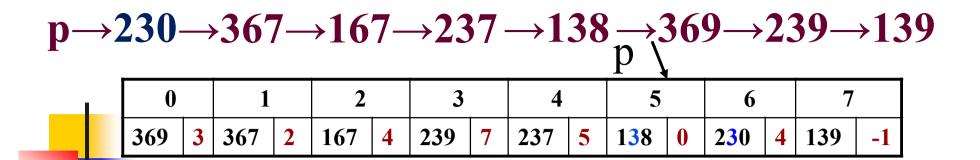
Ρ\														_	
0		1		2		3		4		5		6		7	
369	3	367	2	167	4	239	7	237	5	138	0	230	1	139	-1

$$p \rightarrow 230 \rightarrow 367 \rightarrow 167 \rightarrow 237 \rightarrow 138 \rightarrow 369 \rightarrow 239 \rightarrow 139$$

$p \rightarrow 230 \rightarrow 367 \rightarrow 167 \rightarrow 237 \rightarrow 138 \rightarrow 369 \rightarrow 239 \rightarrow 139$

$p \rightarrow 230 \rightarrow 367 \rightarrow 167 \rightarrow 237 \rightarrow 138 \rightarrow 369 \rightarrow 239 \rightarrow 139$





$p \rightarrow 230 \rightarrow 367 \rightarrow 167 \rightarrow 237 \rightarrow 138 \rightarrow 369 \rightarrow 239 \rightarrow 139$ p

													P			
0		1		2		3		4		5		6		7		
369	3	367	2	167	0	239	7	237	5	138	3	230	4	139	-1	

n\

P \															
0		1		2		3		4		5		6		7	
369	3	367	2	167	0	239	7	237	5	138	3	230	4	139	-1

$$p \rightarrow 230 \rightarrow 237 \rightarrow 138 \rightarrow 239 \rightarrow 139$$

													p \				
0		1		2		3		4		5		6		7			
369	1	367	2	167	0	239	7	237	5	138	3	230	4	139	1		

$$p \rightarrow 230 \rightarrow 237 \rightarrow 138 \rightarrow 239 \rightarrow 139 \rightarrow 367 \rightarrow 167 \rightarrow 369$$

进行第三次分配

67 69

进行第三次分配

第三次收集:

$$p \rightarrow 138 \rightarrow 139 \rightarrow 167 \rightarrow 230 \rightarrow 237 \rightarrow 239 \rightarrow 367 \rightarrow 369$$



- #define MAX_NUM_OF_KEY 8//关键字个数最大值
- #define radix 10//队列个数
- #define MAX_SPACE 1000
- typedef struct {
 Keystype keys[MAX_NUM_OF_KEY];

 int next;
 }SLCell;



- typedef struct {
 SLCell R[MAX_SPACE];
 int keynum;//关键字个数
 int recnum;//待排序数据元素个数
 }SLList;
- typedef int ArrType[radix];



void Distribute(SLCell &R,int i,ArrType &f,Arrtype &r,int head)

```
{ for(j=0;j<radix;j++) f[j]=-1;
    for(p=head;p!=-1;p=R[p].next)
    { j=ord(R[p].keys[i]);
        if (f[j]==-1)f[j]=p;
        else R[r[j]].next=p;
        r[j]=p;
    }
}</pre>
```



void collect(SLCell &R, int i, ArrType f, ArrType r, int &head) { for(j=0; j<Radix &&f[j]==-1; j++);**head**=f[j]; t=r[j]; while(j<Radix) { for (++j; j < Radix-1 & &f[j]==-1; j++);if(f[j]!=-1){ R[t].next=f[j];t=r[j]; } R[t].next=-1;



```
void RadixSort(SLList &L)
{ for(j=0; j<L.recnum-1; j++) L.R[j].next=j+1;
    L.R[L.recnum-1].next=-1;
    head=0;
    for (i=0; i<L.keynum; i++);
    { Distribute (L.R,i,f,r;head); Collect(L.R,i,f,r,head);}</pre>
```

各种排序方法的综合比较

- 时间复杂度--平均的时间性能
- > O(nlogn)--快速排序、堆排序和归并排序
- > O(n²)--直接插入排序、起泡排序和简单选择排序
- > O(n)—基数排序
- 时间复杂度--当待排记录序列按关键字顺序有序时
- > 直接插入排序和起泡排序能达到O(n)的时间复杂度
- > 快速排序的时间性能蜕化为O(n²)
- 简单选择排序、堆排序和归并排序的时间性能不随待 排序数据序列中关键字的分布而改变。

各种排序方法的综合比较

- 空间性能--排序过程中所需的辅助空间大小
- » 所有的简单排序方法(包括:直接插入、起泡和简单选择) 和堆排序的空间复杂度为O(1);
- > 快速排序为O(logn),为递归程序执行过程中, 栈所需的辅助空间
- 》归并排序所需辅助空间最多,其空间复杂度为 O(n);
- > 链式基数排序需附设队列首尾指针,则空间复杂度为 O(rd)。