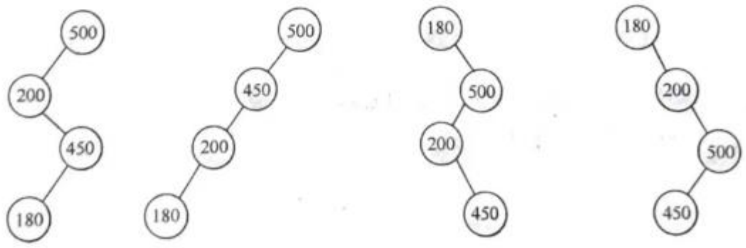
查找

1. A 2015 年考研真题

如下图所示，画出查找路径图，因为折半查找的判定树是一棵二叉排序树，因此看其是否满足二叉排序树的要求。



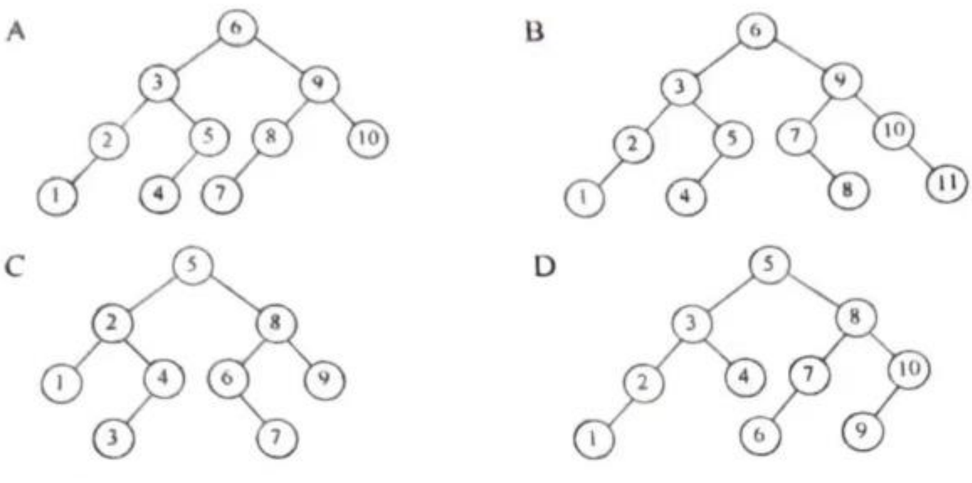
显然，选项 A 的查找路径不满足。

2. C

为使查找效率最高，每个索引块的大小应是 $\sqrt{65025}=255$，为每个块建立索引，则索引表中索引项的个数为 255。若对索引项和索引块内部都采用折半查找，则查找效率最高，为 $\lceil \log_2(255+1) \rceil + \lceil \log_2(255+1) \rceil = 16$。

3. A 2017 年考研真题

折半查找判定树是一棵二叉排序树，其中序序列是一个有序序列。可在树结点上依次填上相应的序号，符合折半查找规则的树即为所求，如下图所示。折半查找算法在选取中间结点时，要么采用向上取整的方式，要么采用向下取整的方式。B 选项，4、5 相加除以 2 采用的是向上取整，7、8 相加除以 2 采用的是向下取整，矛盾。C 选项，3、4 相加除以 2 采用的是向上取整，6、7 相加除以 2 采用的是向下取整，矛盾。D 选项，1、10 相加除以 2 采用的是向下取整，6、7 相加除以 2 采用的是向上取整，矛盾。只有 A 符合折半查找规则。



4. C 2018 年考研真题

根据题意，得到的 HT 如下：

0	1	2	3	4	5	6
	22	43	15			

$ASL_{成功} = (1+2+3)/3 = 2$。

5. 2010年考研真题

1）由装填因子0.7和数据总数7，得一维数组大小为7/0.7 = 10，数组下标为0～9。所构造的散列函数值如下所示：

key	7	8	30	11	18	9	14
H(key)	0	3	6	5	5	6	0

采用线性探测再散列法处理冲突，所构造的散列表为

地址	0	1	2	3	4	5	6	7	8	9
关键字	7	14		8		11	30	18	9	

2）查找成功时，在等概率情况下，查找每个表中元素的概率是相等的。因此，根据表中元素的个数来计算平均查找长度，各关键字的比较次数如下所示：

key	7	8	30	11	18	9	14
次数	1	1	1	1	3	3	2

故 $ASL_{成功}$ = 查找次数/元素个数 = (1 + 2 + 1 + 1 + 1 + 3 + 3)/7 = 12/7。

在计算查找失败时的平均查找长度时，要特别注意防止思维定式，在查找失败的情况下既不是根据表中的元素个数，也不是根据表长来计算平均查找长度的。

查找失败时，在等概率情况下，经过散列函数计算后只可能映射到表中的0～6位置，且映射到0～6中任意一个位置的概率是相等的。因此，是根据散列函数（mod后面的数字）来计算平均查找长度的。在等概率情况下，查找失败的比较次数如下所示：

H(key)	0	1	2	3	4	5	6
次数	3	2	1	2	1	5	4

故 $ASL_{不成功}$ = 查找次数/散列后的地址个数 = (3 + 2 + 1 + 2 + 1 + 5 + 4)/7 = 18/7。

6. 如下

由散列函数可知散列地址的范围为0～10。

采用线性探测法构造散列表时，首先应计算出关键字对应的散列地址，然后检查散列表中对应的地址是否已经有元素。若没有元素，则直接将该关键字放入散列表对应的地址中；若有元素，则采用线性探测的方法查找下一个地址，从而决定该关键字的存放位置。

采用链地址法构造散列表时，在直接计算出关键字对应的散列地址后，将关键字结点插入此散列地址所在的链表。

具体解答如下。

1）线性探测法。

H(1)=1，无冲突，地址1存放关键字1。H(13)=2，无冲突，地址2存放关键字13。H(12)=1，发生冲突，根据线性探测法：H_1=2，发生冲突，继续探测 H_2=3，无冲突，于是12存放在地址为3的表项中。H(34)=1，发生冲突，根据线性探测法：H_1=2，发生冲突，H_2=3，发生冲突，H_3=4，没有冲突，于是34存放在地址为4的表项中。

同理，可以计算其他的数据存放情况，最后结果如下表所示。

散列地址	0	1	2	3	4	5	6	7	8	9	10
关键字	33	1	13	12	34	38	27	22			
冲突次数	0	0	0	2	3	0	1	7			

下面计算平均查找长度：

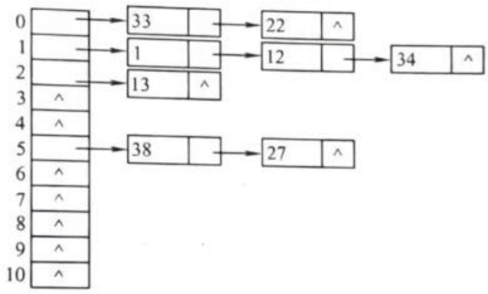
查找成功时，显然查找每个元素的概率都是 1/8。对于 33，由于冲突次数为 0，所以仅需 1 次比较便可查找成功；对于 22，由于计算出的地址为 0，但需要 8 次比较才能查找成功，所以 22 的查找长度为 8；其他元素的分析类似。因此有

$$ASL_{成功} = (1+1+1+3+4+1+2+8)/8 = 21/8$$

查找失败时，由于 H(key)=0~10，因此对每个位置查找的概率都是 1/11，对于计算出的地址为 0 的关键字 key0，只有探测完 0～8 号地址后才能确定该元素不在表中，比较次数为 9；对于计算出的地址为 1 的关键字 key1，只有探测完 1～8 号地址后，才能确定该元素不在表中，比较次数为 8，以此类推。而对于计算出的地址为 8, 9, 10 的关键字，这些单元中没有存放元素，所以只需比较 1 次便可确定查找失败，因此有

$$ASL_{失败} = (9+8+7+6+5+4+3+2+1+1+1)/11 = 47/11$$

2）链地址法构造的表如下：



在链地址表中查找成功时，查找关键字为 33 的记录需进行 1 次比较，查找关键字为 22 的记录需进行 2 次比较，以此类推。因此有

$$ASL_{成功} = (1\times4 + 2\times3 + 3)/8 = 13/8$$

查找失败时，对于地址 0，比较 3 次后确定元素不在表中（空指针算 1 次），所以其查找长度为 3；对于地址 1，其查找长度为 4；对于地址 2，查找长度为 2；以此类推。因此有

$$ASL_{失败} = (3+4+2+1+1+3+1+1+1+1+1)/11 = 19/11$$

值得注意的是，求查找失败的平均查找长度有两种观点：其一，认为比较到空结点才算失败，所以比较次数等于冲突次数加 1；其二，认为只有与关键字的比较才算比较次数。

排序

1．C

快排的阶段性排序结果的特点是，第 i 趟完成时，会有 i 个以上的数出现在它最终将要出现的位置，即它左边的数都比它小，它右边的数都比它大。题目问第二趟排序的结果，即要找不存在两个这样的数的选项。选项 A 中 2, 3, 6, 7, 9 均符合，所以选项 A 排除；选项 B 中，2, 9 均符合，所以选项 B 排除；选项 D 中 5, 9 均符合，所以选项 D 排除；最后看选项 C，只有 9 一个数符合，所以选项 C 不可能是快速排序第二趟的结果。

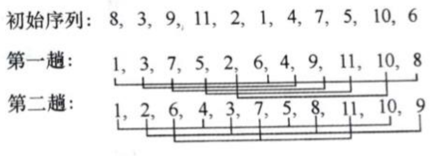
2. C 从后向前“冒泡”的过程为,第 1 趟{11,8,9,10,4,5,6,20,2}，第 2 趟{1,2,8,9,10,4,5,6,20}，第 3 趟{11,2,4,8,9,10,5,6,20},第 4 趟{1,2,4,5,8,9,10,6,20}，第 5 趟{1,2,4,5,6,8,9,10,20}，经过第 5 趟冒泡后，序列已经全局有序，故选择选项 C。实际每趟冒泡发生交换后可以判断是否会导致新的逆序对，如果不会产生，则本趟冒泡之后序列全局有序，所以最少 5 趟即可。
3. B 每趟冒泡和选择排序后，总会有一个元素被放置在最终位置上。显然，这里{11,12}和{4,5}所处的位置并不是最终位置，因此不可能是冒泡和选择排序。2 路归并算法经过第

二趟后应该是每 4 个元素有序的，但{11,12,13,7}并非有序，因此也不可能是 2 路归并排序。

4. A 对于 I，简单选择排序每次选择未排序序列中的最小元素放入其最终位置。对于选项 II，希尔排序每次对划分的子表进行排序，得到局部有序的结果，所以不能保证每趟排序结束都能确定一个元素的最终位置。对于选项 III，快速排序每趟排序结束后都将枢轴元素放到最终位置。对于选项 IV，堆排序属于选择排序，每次都将大根堆的根结点与表尾结点交换，确定其最终位置。对于选项 V，2 路归并排序每趟对子表进行两两归并，从而得到若干局部有序的结果，但无法确定最终位置。
5. D 折半插入排序与直接插入排序都将待插入元素插入前面的有序子表，区别是:确定当前记录在前面有序子表中的位置时，直接插入排序采用顺序查找法，而折半插入排序采用折半查找法。排序的总趟数取决于元素个数 n，两者都是 n-1 趟。元素的移动次数都取决于初始序列，两者相同。使用辅助空间的数量也都是 O(1)。折半插入排序的比较次数与序列初态无关，为 O (nlogn)；而直接插入排序的比较次数与序列初态有关，为 $O(n)$~$O(n^2)$。
6. D 2018 年考研真题

第一趟分组：8, 1, 6；3, 4；9, 7；11, 5；2, 10；间隔为 5，排序后组内递增。

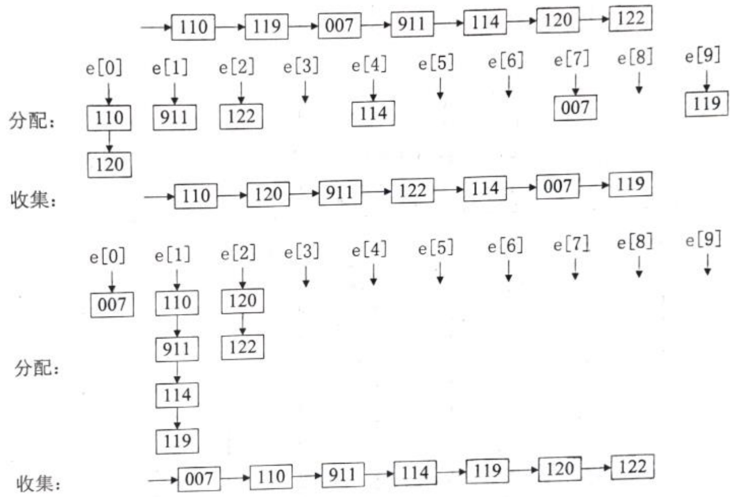
第二趟分组：1, 5, 4, 10；3, 2, 9, 8；7, 6, 11；间隔为 3，排序后组内递增。



故答案选 D。

7. C 2013 年考研真题

基数排序的第 1 趟排序是按照个位数字的大小来进行的，第 2 趟排序是按照十位数字的大小来进行的，排序的过程如下图所示。



8. C 基数排序是一种稳定的排序方法。由于采用最低位优先(LSD)的基数排序，即第一趟对个位进行分配和收集操作，因此第一趟分配和收集后的结果是{151,301,372,892,93,43,485,946,146,236,327,9}，元素 372 之前、之后紧邻的元素分别是 301 和 892。
9. D 2022 年考研真题

直接插入排序和快速排序的特点如下表所示。

	适合初始序列情况	适合元素数量	空间复杂度	稳定性
直接插入排序	大部分元素有序	较少	$O(1)$	稳定
快速排序	基本无序	较多	$O(\log_2 n)$	不稳定

可见，选项 I、II、III、IV 都是采用直接插入排序而不采用快速排序的可能原因。

10. 2016 年考研真题，核心思想就是变形的快排

由题意知，将最小的$\lfloor n/2 \rfloor$个元素放在 A_1 中，其余的元素放在 A_2 中，分组结果即可满足题目要求。仿照快速排序的思想，基于枢轴将 n 个整数划分为两个子集。根据划分后枢轴所处的位置 i 分别处理：

① 若 $i = \lfloor n/2 \rfloor$，则分组完成，算法结束。

② 若 $i < \lfloor n/2 \rfloor$，则枢轴及之前的所有元素均属于 A_1，继续对 i 之后的元素进行划分。

③ 若 $i > \lfloor n/2 \rfloor$，则枢轴及之后的所有元素均属于 A_2，继续对 i 之前的元素进行划分。

基于该设计思想实现的算法，无须对全部元素进行全排序，其平均时间复杂度是 $O(n)$，空间复杂度是 $O(1)$。

思考：代码该怎么写？