

DSA真题2014 + 解答

一、判断

1. 由同一组共 n 个词条构成的任意两棵BST，经 $O(\log n)$ 次zig或zag旋转之后，必定可以相互转换。
– 错误。详见习题解析7-15，至多需要经过 $2n-2$ 次旋转调整，是 $O(n)$ 级别的。
2. 设在有向图 G 中，存在一条自顶点 v 通往 u 的路径。于是，若在某次DFS中有 $dTime(v) < dTime(u)$ ，则这次DFS所生成的DFS森林中， v 必定是 u 的祖先。
– 错误。构造的反例具体看16大题第一道。
3. 在无向连通图 G 中选定一个顶点 s ，并将各顶点 v 到 s 的距离记作 $dist(v)$ （特别地， $dist(s)=0$ ）。于是在 $G.Bfs(s)$ 过程中，若辅助队列为 Q ，则 $dist(Q.front()) + 1 \geq dist(Q.rear())$ 始终成立。
– 正确。任意时刻在 Q 内， $dist$ 不会相差超过1。
4. 我们知道，因同一顶点的邻居被枚举的次序不同，同一有向图 G 所对应的DFS森林未必唯一。然而只要起始于 G 中某顶点 s 的某次DFS所生成的是一棵树，则起始于 s 的任何一次DFS都将生成一棵树。
– 正确。记住DFS为遍历过程，所有从 s 能够到达的都会被记录；又因为已存在一棵树，故 s 可达其余所有点。
5. 无论 g 和 h 互素与否，已经 h -有序的序列再经 g -排序之后，必然继续保持 h -有序。
– 正确。
6. 在BM算法中，对于任一模式串 P ， $0 < gs(j) \leq j$ 对于每个 $0 \leq j < |P|$ 都成立。
– 正确。
7. 相对于二叉堆，尽管多叉堆的高度更低，但无论是下滤一层还是整个下滤过程，时间成本反而都会增加。
– 错误。对于三叉堆来说，整个下滤过程的时间成本下降了。
8. 我们知道，采取双向平方试探策略时，应该将散列表取作素数 $M = 4k + 3$ 。尽管这样可以极大降低查找链前 M 个位置发生冲突的概率，但仍不能杜绝。
– 错误。在前 M 个可以避免冲突的发生。
9. 即便访问序列不满足局部性（比如完全理想的随机），伸展树依然能够保证分摊 $O(\log n)$ 的性能。
– 正确。习题解析8-2有证明；伸展树所有基本操作接口的分摊时间复杂度均为 $O(\log n)$ 。

10. 我们知道, `BTree:solveOverflow()`和`BTree:solveUnderflow()`在最坏情况下均需下界($\log n$)的时间, 然而在B-树任一足够长的生命期内, 就分摊意义而言二者都仅需要 $O(1)$ 时间。
- 正确。习题解析8-6、8-7; 根据结论得到, 分裂与合并的总次数是 $O(n)$ 级别的, 分摊下来就是 $O(1)$ 时间的。

二、选择

1. 设在某新节点插入AVL树后(尚待平衡化时), 最低失衡节点为g。若此时g的左、右孩子的平衡因子分别为-1, 0, 则应通过 (C) 旋转使之重新恢复平衡。
- A) zig;
 - B) zig+zag;
 - **C) zag+zig;**
 - D) zag;
 - E) 不确定
2. 为从2014个随机元素中挑选出最大的5个, (A) 在最坏情况下所需的比较操作次数最少。
- **A) 构建大顶的锦标赛树, 再做5次delMax();**
 - B) 用Floyd算法构建大顶堆, 再做5次delMax();
 - C) 采用选择排序算法, 但仅执行前5次迭代;
 - D) 采用起泡排序算法, 但仅执行前5次迭代;
 - E) 用linearSelect()算法找出第5大的元素, 再遍历一趟找出(至多)4个大于它的元素
3. 有向图的DFS不仅在起点任意, 而且每一步迭代往往都会有多个顶点可供选择, 故所生成的**DFS森林**并不唯一确定, 且其中所含 (ABCD) 的数量也可能不同。
- **A) 树边**
 - **B) 前向边**
 - **C) 后向边**
 - **D) 跨越边**
 - E) 以上皆非
4. 相对于除余法, MAD法在 (BC) 方面有所改进。
- A) 计算速度
 - **B) 高阶均匀性**
 - **C) 不动点**
 - D) 满射性
 - E) 以上皆非
5. 将硬币换成理想的骰子, 且约定投出“6”时新塔才停止生长。于是对于同样存放n个元素的跳转表而言, (BCD) 的期望值将有所增长, 但仍保持 $O(1)$ 。
- A) 查找过程中, 在同一高度连续跳转的次数
 - **B) 查找过程中, 由“向右”到“向下”转折的次数**
 - **C) 查找过程中, 沿同一座塔连续下行的层数**

- **D)** （在查找定位之后）为创建一座新塔所需的时间
6. （ BEF ）属于针对闭散列策略的冲突排解方法。
- A) multiple slots
 - **B) linear probing**
 - C) overflow area
 - D) separate chaining
 - **E) quadratic probing**
 - **F) double hashing**
 - 注：ACD都是针对开散列的
7. 对于任何一颗二叉树T，其右、左子树的规模之比“ $\lambda = T.re().size() / T.le().size()$ ”称作右偏率。对于（常规）高度同为h的AVL树（A），红黑树（R），左式堆（L），若分别考察其 λ 所能达到的最大值，则在h足够大之后，三者按此指标的排列次序应是（ B ）。
- A) $L < R < A$
 - **B) $L < A < R$**
 - C) $R < A < L$
 - D) $A < R < L$
 - E) 以上皆非
 - 注：左式堆的节点往左偏，所以右偏率应该最小；而红黑树相比于AVL更加宽松，所以红黑树最大值肯定要比AVL要大
8. 以下数据结构中，空间复杂度不超过线性的有（ ABDEF ）。
- **A) 2d-tree**
 - **B) 3d-tree**
 - C) 2D range tree
 - **D) interval tree**
 - **E) segment tree**
 - **F) priority search tree**
 - 注：C是 $n \log n$ 空间的，注意segment tree书上的空间复杂度写错了

三、综合

具体的题目请看2014秋·样题，考试题中题型不变，但是数据有些许改变。