通过hdu2665来理解可持久化线段树

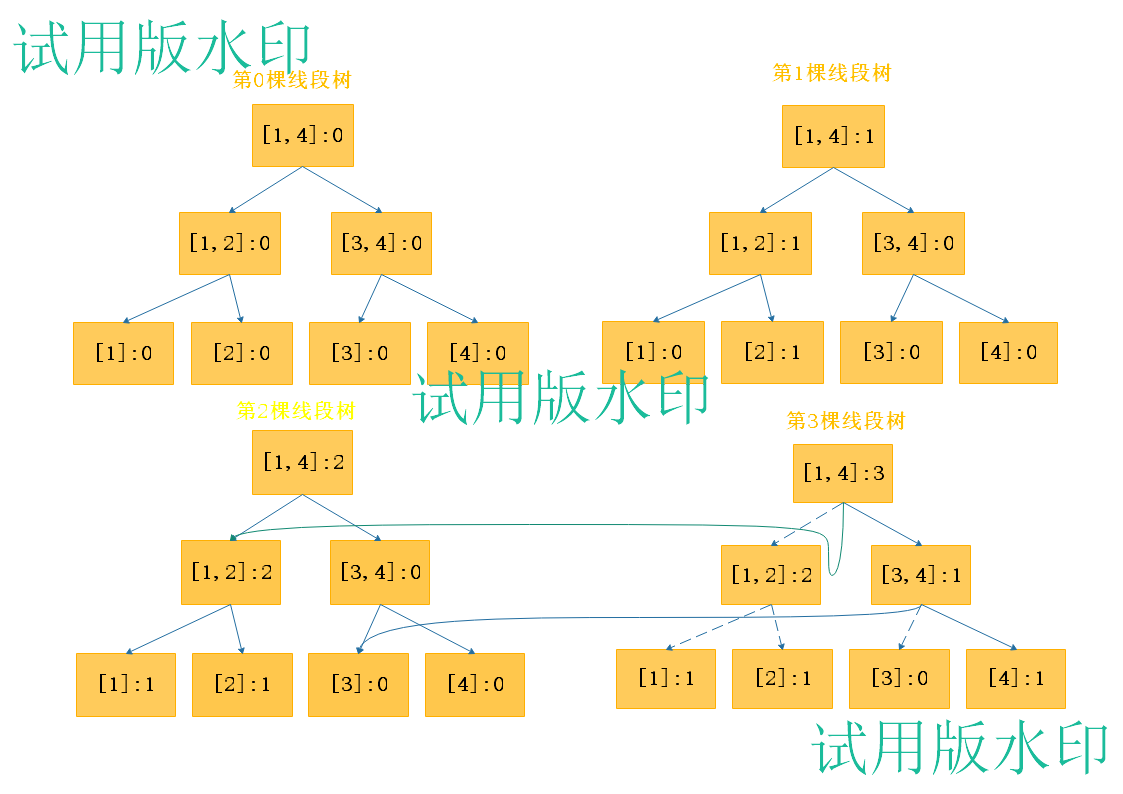
求区间第k大的数

什么是可持久化数据结构

可持久化数据结构(Persistent data structure)就是利用函数式编程的思想使其支持询问历史版本、同时充分利用它们之间的共同数据来减少时间和空间消耗。

在这里我们可以建立n颗线段树，第i颗线段树表示对区间[1,i]的所有data[i]建立线段树，而线段树的区间范围是区间[1,m]的取值范围。因为data[i]的数据可能很大，我们先对它进行离散化处理。下面举个例子。

有个数列{4,3,10,7}，离散化之后3🡪1,4🡪2,10🡪4,7🡪3。



我们怎么求区间[x,y]的第k大？

这时我们考虑第x-1棵线段树，和第y棵线段树。两颗线段树在根节点[1,m]的差值就是区间[x,y]的元素个数。Res = Tr[y].lc.cnt-Tr[x-1].rc.cnt > k,则说明第k大在区间[1,mid]内，否则在[mid+1,m]区间内，这是我们在这个区间内找第k-res大的数。

这样建n颗线段树，肯定会MLE的，我们观察下第i，i+1棵线段树。在第i颗线段树的基础上插入data[i+1]就得到了第i+1棵线段树，所以第i+1棵线段树与第i棵线段树只有logn个结点不同，所以我们只要添加logn个结点，其他节点与第i棵线段树公用即可。这样的空间复杂度是n\*4+nlogm（m是线段树区间范围）。

时间复杂度是2nlogm