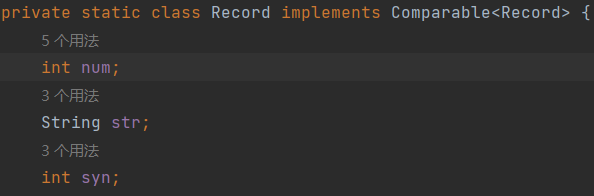
**实验二、数据库系统算法的实现**

姓名：瞿久尧 学号：120L022314

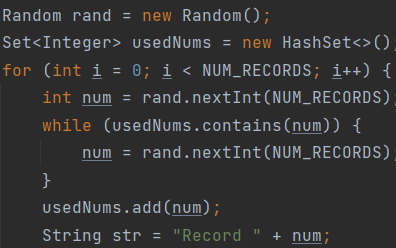
1. 实验环境的构建
   1. **记录生成**

每条记录使用一个类对象Record来记录，Record中包含3个属性：num为数值型，str为字符串型，syn为int型，并不参与记录，只是会在临时文件中记录，表示该record所在块号

同时该类中会重写compareTo方法，方便后期直接调用集合的排序方法



记录为随机生成，但需要保证20w条记录不重复，因此需要用集合来记录已经存在的记录



* 1. **文件准备**

使用FileOutputStream和文件名data.dat创建文件输出流用于存储我们随机生成的20w条记录

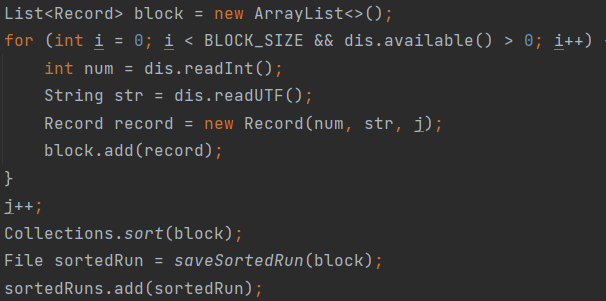


将记录中的num和str写入文件中

1. 基本算法设计
   1. **分块并对块中记录进行排序**

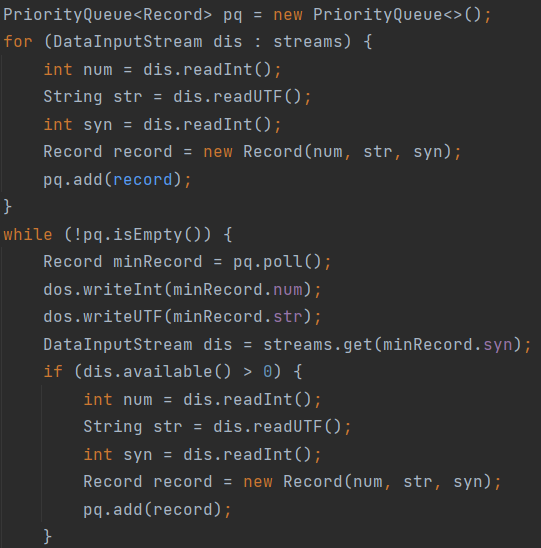
由于记录太多，因此选择将20w条记录分成20份，每块1w条记录，每读取1块，就将这一块进行排序并存到一个临时文件中

这里排序采用的是快速排序，时间复杂度为O(nlogn)

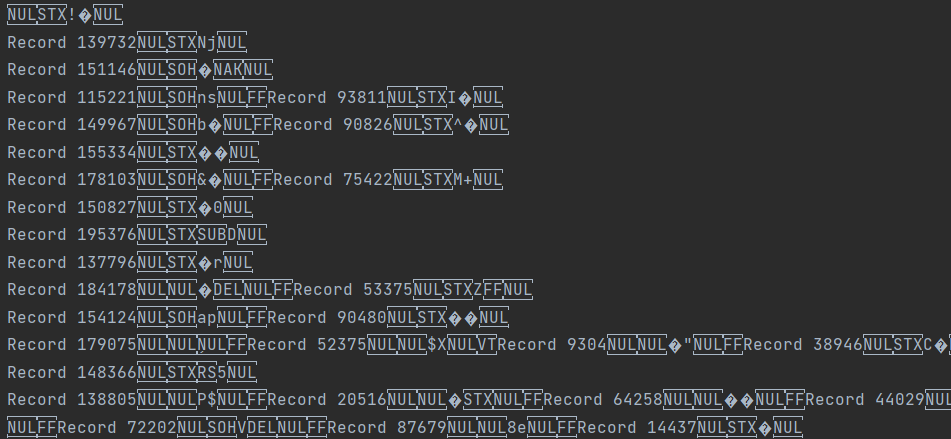


* 1. **对硬盘中的文件进行归并排序**

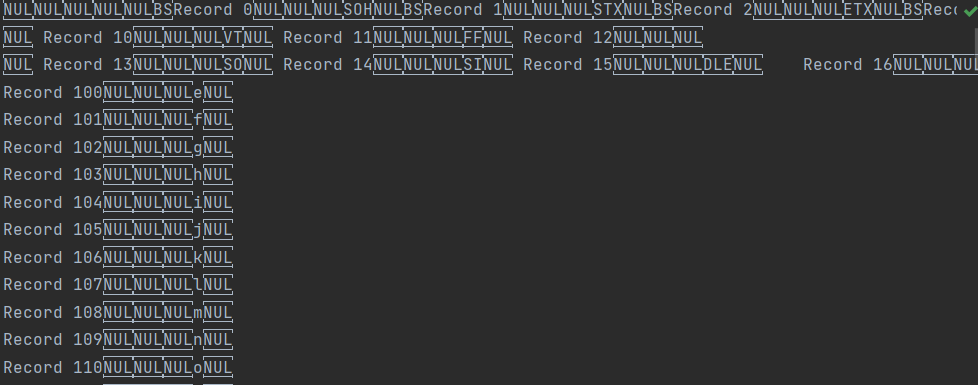
将存在硬盘中的临时文件进行归并排序，这里使用了优先队列的数据结构，将每个流的第一条数据放入优先队列中，然后从优先队列中挑选最小的数据写到目标文件中，随后根据被挑选出来的数据的块号syn找到对应的输入流，在对应输入流加入下一条数据到优先队列中，直到优先队列所有记录都被写入到目标文件中



1. 程序运行结果
   1. **随机生成记录**



* 1. **排序后的记录**



* 1. **输出总时间**

