



外部排序 | 外部排序模型

外排序

大多数内排序算法都是利用了内存是直接访问的事实,读写一个数据是常量的时间。如果输入是在磁带上,磁带上的元素只能顺序访问。甚至数据是在磁盘上,效率还是下降,因为转动磁盘和移动磁头会产生延迟。

- ④ 外排序模型

- ④ 预处理

- ④ 归并

外排序模型

- ④ 外排序具有设备依赖性。这里考虑的算法工作在磁带上
- ④ 完成有效的排序至少需要两个磁带机
- ④ 三个磁带机可以简化问题

外排序的基本方法

- ◎ 由于一次外存操作所需的时间可以执行数百条甚至上千条指令，因此在外排序中我们主要考虑的是如何**减少外存储器的读写**
- ◎ 在外存上进行排序的最常用的方法是利用**归并排序**，因为归并排序只需要访问被归并序列中的第一个元素，这非常**适合于顺序文件**。
- ◎ 外排序由两个阶段组成：
 - 预处理阶段：根据内存的大小将一个有 n 个记录的文件分批读入内存，用各种内排序算法排序，形成一个个有序片段。
 - 归并阶段：将这些有序片段逐步归并成一个有序文件。