

1. 假设一个记录长为 32 个字节，一个块长 1024 个字节（每个块有 32 个记录），工作内存是 1MB（还有用于 I/O 缓冲区、程序变量等的其他存储空间）。使用置换选择和多路归并，其中归并算法只允许扫描两遍。预计能得到的文件最长为多少？解释这个结果是怎样计算出来的。

工作内存可个储存 1MB 内容

因此置换选择所得到的顺串平均大小为 2MB

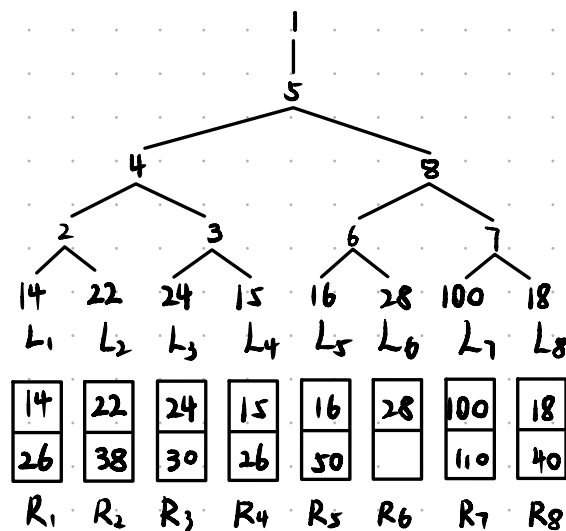
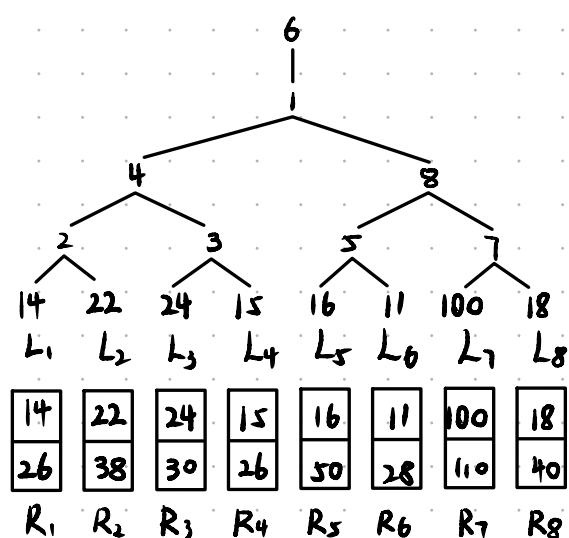
对外存的访问以 1KB 的块为连续单位，这 1KB 在同一个顺串中

因此工作内存最多同时存  $\frac{1MB}{1KB} = 1024$  个不同的顺串的数据

从而最多为 1024 路归并

在 2 遍归并后平均最大长度为  $2MB \times 2^{10} \times 2^{10} = 2TB$

2. 现有 8 个顺串，每个顺串的第一个记录的关键码分别为 14, 22, 24, 15, 16, 11, 100, 18, 第二个记录的关键码分别为 26, 38, 30, 26, 50, 28, 110, 40. 请画出对顺串开始 8 路归并时的败者树，以及输出一个优胜者后重构得到的败者树。

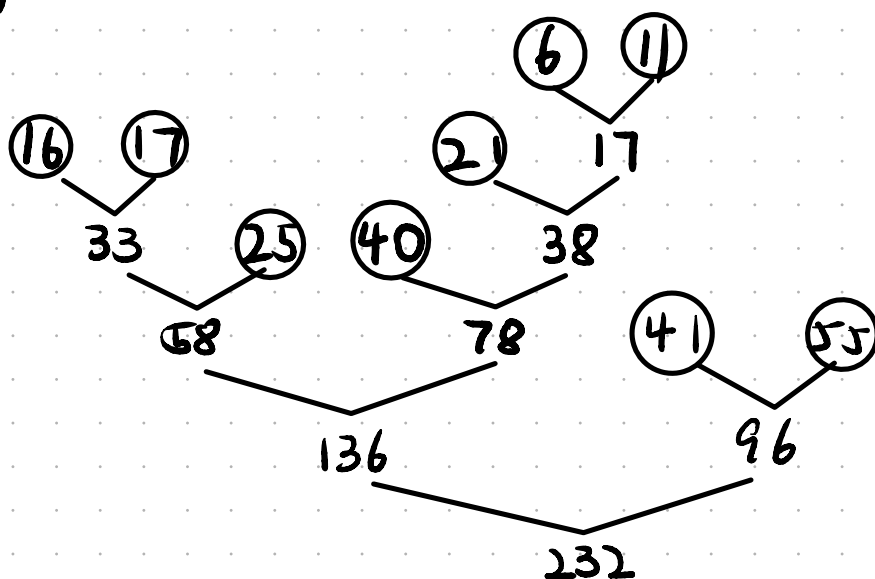


3. 现在有 9 个长度不同的的顺串，其长度分别为 17, 40, 16, 55, 25, 11, 6, 21, 41。请用二路归并的方法对其进行归并。

(1) 构造最佳归并树。

(2) 根据最佳归并树计算每一趟和总的读记录次数。

(1)



(2)  $(6+11) \times 5 + (16+17+21) \times 4 + (25+40) \times 3 + (41+55) \times 2$   
 $= 688$