

第 12 章 文件

教材中练习题及参考答案

1. 什么是文件的逻辑记录和物理记录？它们有什么区别与联系？

答：记录是文件存取操作的基本单位。逻辑记录是按用户观点描述的基本存取单位，物理记录是按外存设备观点描述的基本存取单位。

通常逻辑记录和物理记录之间存在三种关系：

- (1) 一个物理记录存放一个逻辑记录。
- (2) 一个物理记录包含多个逻辑记录。
- (3) 多个物理记录表示一个逻辑记录。

2. 比较顺序文件、索引非顺序文件、索引顺序文件和哈希文件的存储代价、检索、插入及删除记录时的优点和缺点。

答：(1) 顺序文件只能按顺序查找法存取，即按记录的主关键字逐个查找。这种查找法对于少量的检索是不经济的，但适合于批量检索。顺序文件的存取优点是速度快。顺序文件不能像顺序表那样的方法进行插入，删除和修改，因为文件中的记录不能像数组空间的数据那样“移动”，而只能通过复制整个文件的方法来实现上述更新操作。

(2) 索引非顺序文件适合于随机存取，这是由于主文件的记录是未按关键字排序的，若要进行顺序存取将会频繁地引起磁头移动，因此索引非顺序文件不适合于顺序存取。在索引顺序文件中，由于主文件也是有序的，所以它既适合于直接存取，也适合于顺序存取。另一方面，索引非顺序文件的索引是稠密索引，而索引顺序文件是稀疏索引，故后者的索引减少了索引项的数目，虽然它不能进行“预查找”，但由于索引占用空间较少、管理要求低，因而提高了索引表的查找速度。

(3) 哈希文件也称为直接存取文件，是利用哈希法组织文件，它类似于哈希表，根据文件中关键字的哈希函数值和处理冲突的方法，将记录哈希到外存上。这种文件组织方法只适用于像磁盘这样的直接存取设备。哈希文件的优点是：文件可以随机存放，记录不必进行排序，插入、删除操作方便，存取速度快，无需索引区，因而节省存储空间。哈希文件的缺点是：不能进行顺序存取且访问方式也只限于简单查询，此外在经过多次插入、删除后可能出现文件结构不合理以及记录分布不均匀等现象，这时需要重组文件，但这个工作是很费时的。

3. 某职工表文件如表 12.1 所示，其中以职工号为主关键字。

- (1) 若将职工文件组织成顺序有序文件，请给出文件的存储结构。
- (2) 若将该文件组织成索引非顺序文件，请给出其索引表结构。

(3) 若将该文件组织成多重表文件, 请给出主文件结构及性别索引, 工作时间索引(只考虑年份)及年龄段(10岁为一个年龄段)索引。

(4) 若将该文件组织成倒排文件, 请写出性别索引、职称索引及年龄段索引, 组织索引要求同(3)。

答: (1) 该文件若组织为顺序有序文件, 则应将职工文件记录按职工号有序存放, 每个记录信息不变。假设其顺序有序文件仍存储在原文件区域, 则它的文件结构如表 12.2 所示(假设物理地址编号从 1 开始)。

表 12.1 职工表

物理地址	职工号	姓名	年龄	性别	工作时间	职称
1	105	李华	36	男	1997.7	副教授
2	125	王丽	42	女	1984.7	副教授
3	108	张英	28	男	1998.7	讲师
4	182	陈军	52	男	1970.9	教授
5	135	吴斌	25	男	1999.9	助教
6	116	章萍	58	女	1965.9	教授
7	140	华明	40	男	1989.7	讲师

(2) 若该文件组织成索引非顺序文件, 则主文件即为表 12.1 的职工文件, 其索引文件如表 12.3 所示(假设物理地址编号从 11 开始)。

(3) 若将该文件组成多重链表文件, 其主文件如表 12.4 所示; 其性别索引、职称索引和年龄索引分别如表 12.5、表 12.6、表 12.7 所示。

(4) 若将该文件组织成倒排文件, 主文件如表 12.1 所示, 对应的性别索引、职称索引及年龄段索引分别如表 12.8、表 12.9 和表 12.10 所示。

表 12.2 职工的顺序有序文件存储结构

物理地址	职工号	姓名	年龄	性别	工作时间	职称
1	105	李华	36	男	1997.7	副教授
2	108	张英	28	男	1998.7	讲师
3	116	章萍	58	女	1965.9	教授
4	125	王丽	42	女	1984.7	副教授
5	135	吴斌	25	男	1999.9	助教
6	140	华明	40	男	1989.7	讲师
7	182	陈军	52	男	1970.9	教授

表 12.3 索引表

物理地址	职工号	物理地址
11	105	1
12	108	3
13	116	6
14	125	2
15	135	5
16	140	7
17	182	4

表 12.4 多重表主文件

物理地址	职工号	姓名	年龄	性别	工作时间	职称	性别链	年龄链	职称链
1	105	李华	36	男	1997.7	副教授	3	^	2
2	125	王丽	42	女	1984.7	副教授	6	7	^
3	108	张英	28	男	1998.7	讲师	4	5	7
4	182	陈军	52	男	1970.9	教授	5	6	6
5	135	吴斌	25	男	1999.9	助教	7	^	^
6	116	章萍	58	女	1965.9	教授	^	^	^
7	140	华明	40	男	1989.7	讲师	^	^	^

表 12.5 性别索引

次关键字	头指针	链长
男	1	5
女	2	2

表 12.6 职称索引

次关键字	头指针	链长
助教	5	1
讲师	3	2
副教授	1	2
教授	4	2

表 12.7 年龄索引

次关键字	头指针	链长
20-29	3	2
30-39	1	1
40-49	2	2
50-59	4	2

表 12.8 性别倒排索引

次关键字	物理地址
男	1, 3, 4, 5, 7
女	2, 6

表 12.9 称倒排索引

次关键字	物理地址
助教	5
讲师	3, 7
副教授	1, 2
教授	4, 6

表 12.10 职称倒排索引

次关键字	物理地址
20~29	3, 5
30~39	1
40~49	2, 7
50~59	4, 6