

第8章

1. 简述数据库堆文件的链表和页目录两种组织方式中，空闲空间分别是如何组织的？

答：链表方式下，当只有一个堆文件时，堆文件头部设立一个**header** 页面，并存放两个指针，其中一个指针指向空页列表（**free list**）头部，而在每个页面内部，均记录当前空闲的空间（**slot**）；当有多个堆文件时，需要元数据记录文件中有哪些页面，以及哪些页有空闲空间。

页目录方式下，堆文件中设立一类专门的页面（目录页），用于记录所有的数据页的存放位置，该目录也同时记录每个页面的空闲空间信息（**slot**）。**DBMS**必须保持目录页与所有页的当前信息同步。

第8章

2. 简述面向元组型的页设计中槽页方案的基本思想，并简要说明槽页方案相对于数组式的元组存储方案的好处。

答：每个页面都有页头，包含页面**ID**，还会保存页面内的**Slot**数组，记录已占用的“槽位（**slot**）”、这些“槽位（**slot**）”映射到特定元组开始位置的偏移量、以及上一次使用槽位的开始位置。

元组的存放则在页内倒序从页尾向页头依次存放。元组在内部的唯一标识符：可以使用页面**ID**和**slot id**（或偏移量），也可包含文件位置信息。

槽页方案元组避免了删除元组时会产生碎片的问题，同时能更好的应对变长元组的存储。

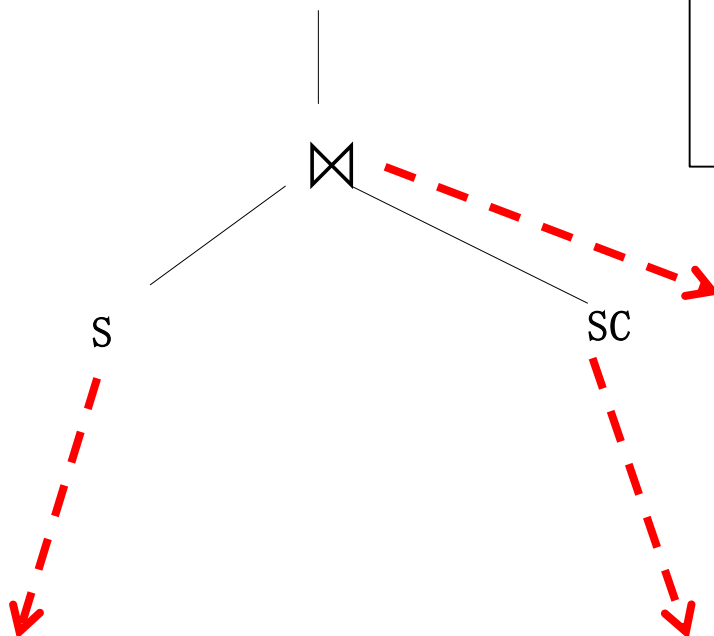
第8章

3. 设有学生关系S(SNO, SNAME, SDEPT)存储学生的学号、姓名和所属院系，学生选课关系SC(SNO, CNO, SCORE)存储学生选课记录的学号、课号和成绩，现有查询的SQL语句“SELECT S.SNO, CNO, SCORE FROM S, SC WHERE S.SNO=SC.SNO”，请画出查询计划的语法树，请写出该查询语法树的处理模型采用向量模型时每个查询算子的伪代码（设算子中每次处理的一个batch的大小为500）。

第8章

3. 答:

$\Pi_{S.SNO,CNO,SCORE}$ 



```
out=[]  
for t in child.Next():  
    out.add(projection(t))  
if |out|>=500: emit(out)
```

```
out=[]  
for S in left.next():  
    buildHashTable(S)  
for SC in right.next():  
    if Probe(SC) out.add(S ⋈ SC)  
if |out|>=500: emit(out)
```

```
out=[]  
for t in S:  
    out.add(t)  
if |out|>=500: emit(out)
```

```
out=[]  
for t in SC:  
    out.add(t)  
if |out|>=500: emit(out)
```

注：虚线不属于语法树结构的内容