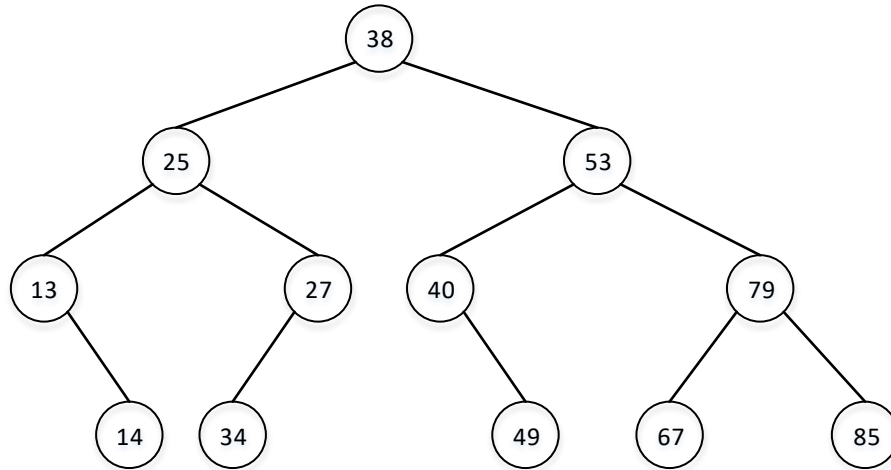


第五章

1、(1) 判定树



(2) 查找 49，依次比较：38,53,40,49

(3) 查找 80，依次比较：38,53,79,85

(4) 每个元素的查找长度为该结点在判定树中的层数，如 40 的查找长度为 3，所以，查找成功平均查找长度： $(3+4+2+3+4+1+3+4+2+4+3+4) / 12 = 3.08$

2、分析：最理想的分块方式应该是使得平均查找长度最小。根据 p206 公式：

$$ASL \approx \log_2 \left(\frac{n}{m} + 1 \right) - 1 + \frac{m+1}{2}$$

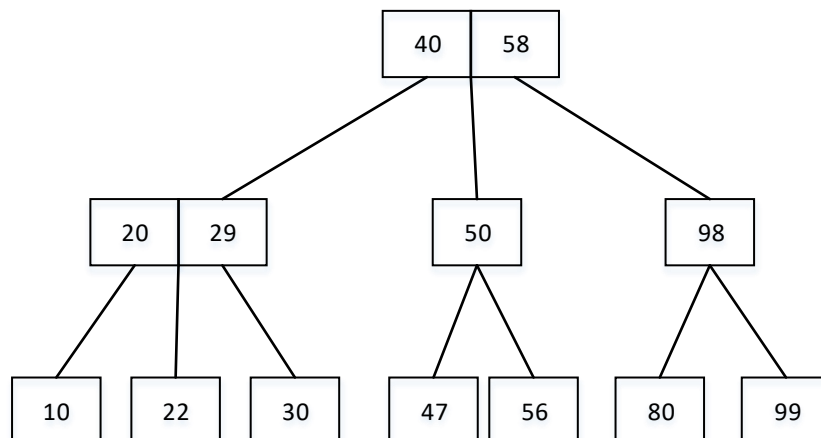
这里 $n=4000$ ，求 m （块内记录数）的值（ $m \in [1,4000]$ ）使得 ASL 取得最小值。

求导可得（或者编写计算机程序求得结果）当 $m=5$ 时，ASL 最小，为 7.15，即分 800 块，每块长度为 5 时最理想。

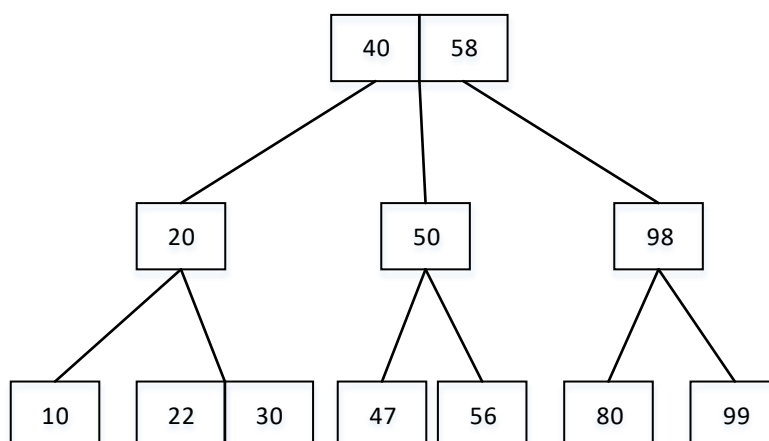
若每块长度为 50，带入上述公式，得出平均查找长度为 27.12

3、

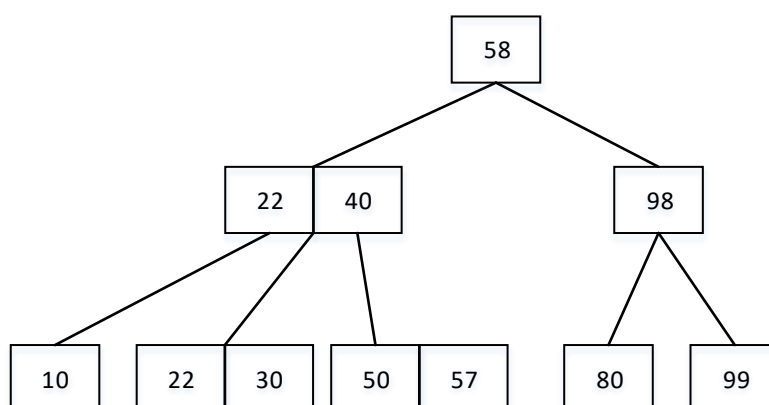
B-树：



删除 29 之后：



删除 47 之后:



4、简述 B-树与 B+树的区别。

- (1) 同阶 B+树结点最大关键字数比 B-树多 1
- (2) B+树所有关键字均在叶子结点，而 B-树关键字不止存在叶子结点
- (3) B+树叶子结点相互连接为双向链表，而 B-树叶结点之间没有直接联系
- (4) B+树非叶结点的指针 $p[i]$ 指向关键字区间在 $[k[i], k[i+1])$ 之间的子树，而 B-树为全开区间

5、(1) 线性探查法:

| | | | | | | | | | | |
|---|---|---|----|----|----|----|----|----|----|----|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| | 4 | | 12 | 49 | 38 | 13 | 24 | 32 | 21 | |

hash(32) = 8

hash(13) = 6

hash(49) = 4

hash(24) = 6 冲突，选择 7

hash(38) = 4 冲突，选择 5

hash(21) = 8 冲突，选择 9

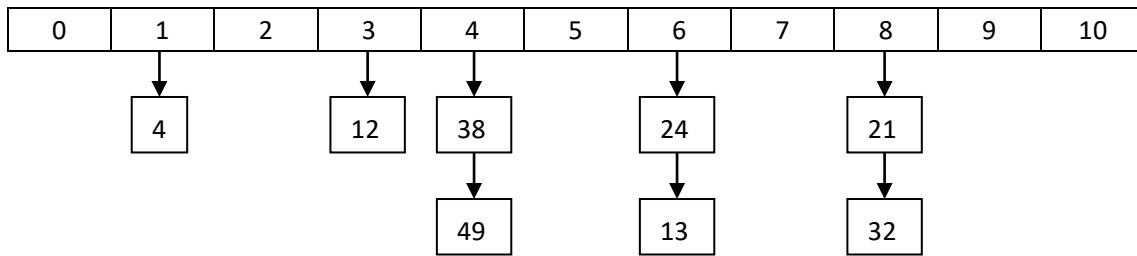
hash(4) = 1

hash(12) = 3

查找成功平均长度: $(1+1+1+2+2+2+1+1)/8 = 11/8$

查找失败平均长度: $(3+2+8+7+6+5+4+3+2)/11 = 40/11$

(2) 链接法:



查找成功平均长度: $(1+1+1+2+1+2+1+2)/8 = 11/8$

查找失败平均长度: $(6+2+2+3+3+3)/11 = 19/11$

6、

插入关键码算法:

设散列表大小为 s , 关键码得到的 hash 地址为 x

从 x 位置开始依次查看 $(x+i)\%s$ ($i=0,1,\dots,s-1$) 位置, 若当前位置不包含有效关键码 (为空或者已删除), 则将关键码存入, 插入成功, 返回; 否则, 插入失败, 返回。

删除关键码算法:

设散列表大小为 s , 关键码为 k , 关键码得到的 hash 地址为 x

从 x 位置开始依次查看 $(x+i)\%s$ ($i=0,1,\dots,s-1$) 位置, 若当前位置包含有效关键码且关键码为 k , 删除关键码并将当前位置标记为已删除; 否则删除失败, 返回。

7、(1) 线性探查法:

| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|----|----|---|----|----|----|---|---|---|----|
| 89 | 71 | | 23 | 73 | 44 | | | | 99 |

hash(71) = 1

hash(23) = 3

hash(73) = 3 冲突, 选择 4

hash(99) = 9

hash(44) = 4 冲突, 选择 5

hash(89) = 9 冲突, 选择 0

查找成功平均查找长度: $(1+1+2+1+2+2)/6 = 3/2$

查找失败平均查找长度: $(4+3+2+4+3+2+4)/10 = 11/5$

(2) 双散列:

| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|----|---|----|----|---|---|----|---|----|
| | 71 | | 23 | 44 | | | 73 | | 99 |

hash(71) = 1

hash(23) = 3

hash(73) = 3 冲突

rehash(73) = 4

$H1 = (\text{hash}(73) + 1 * \text{rehash}(73)) \% 10 = 7$

hash(99) = 9

hash(44) = 4

hash(89) = 9 冲突

rehash(89) = 5

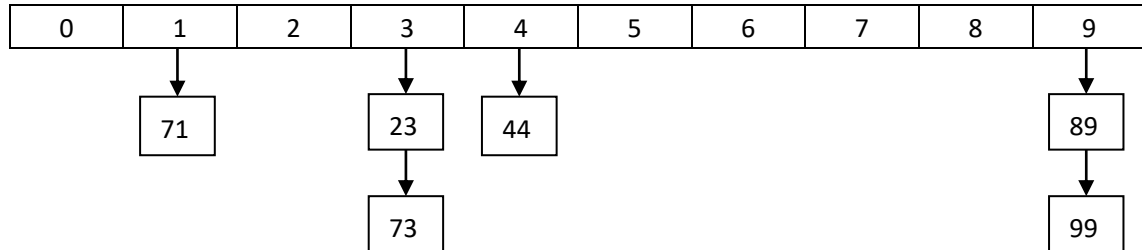
$H1 = (\text{hash}(89) + 1 * \text{rehash}(89)) \% 10 = 4$ 冲突

$H2 = (\text{hash}(89) + 2 * \text{rehash}(89)) \% 10 = 9$ 冲突，且回到 $\text{hash}(89)$ 位置，插入失败

查找成功平均查找长度： $(1+1+2+1+1) / 5 = 6/5$

查找失败平均查找长度：略

(3) 链接法：



查找成功平均查找长度： $(1+1+2+1+1+2)/6 = 4/3$

查找失败平均查找长度： $(6+2+3+2+3)/10 = 16/10$

8、

线性探查再散列的查找成功的平均查找长度的理论值为 $(1 + 1/(1-a))/2$ ，因为比较次数小于 3，得到装填因子 $a \leq 4/5$ ，为 0.8，因此存储空间约为 $250/0.8 = 312.5$ ，上取整为 313。