



数据结构与算法（十一）

张铭 主讲

采用教材：张铭，王腾蛟，赵海燕 编写
高等教育出版社，2008. 6（“十一五”国家级规划教材）

<http://www.jpk.pku.edu.cn/pkujpk/course/sjjg>

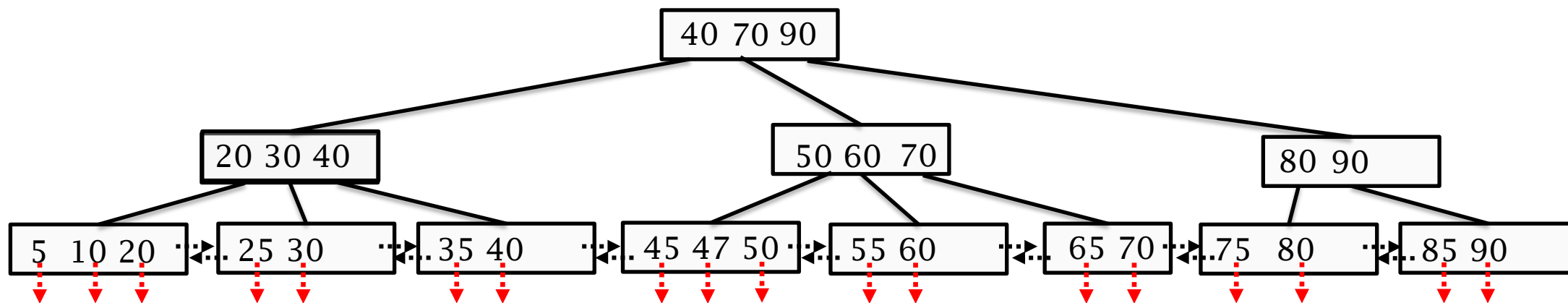


主要内容

- 11.1 线性索引
- 11.2 静态索引
- 11.3 倒排索引
- 11.4 动态索引
 - 11.4.1 B 树
 - 11.4.2 B 树的性能分析
 - 11.4.3 B+ 树
 - 11.4.4 B 树、B+ 树索引性能的比较
- 11.5 位索引技术
- 11.6 红黑树

11.4.3 B⁺ 树

- 是B 树的一种变形，在叶结点上存储信息
 - 所有的关键码均出现在叶结点上
 - 各层结点中的关键码均是下一层相应结点中最大关键码（或最小关键码）的复写

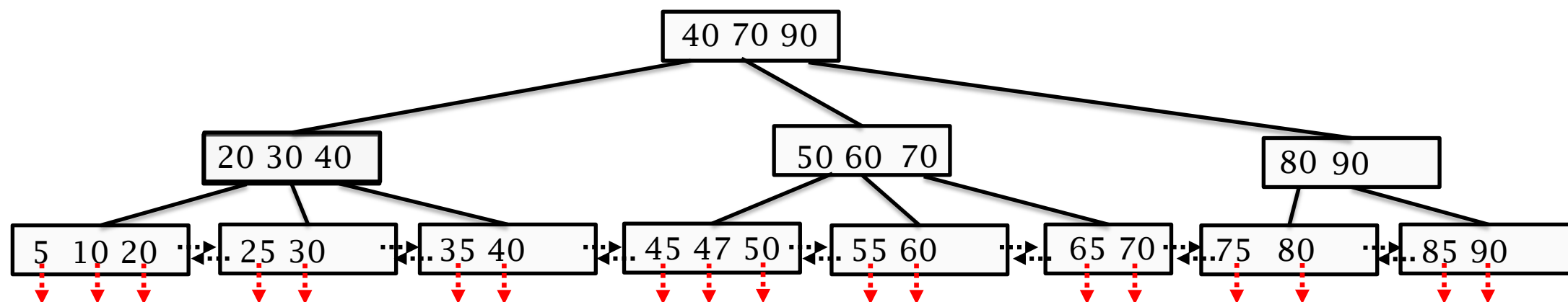




B⁺ 树的结构定义

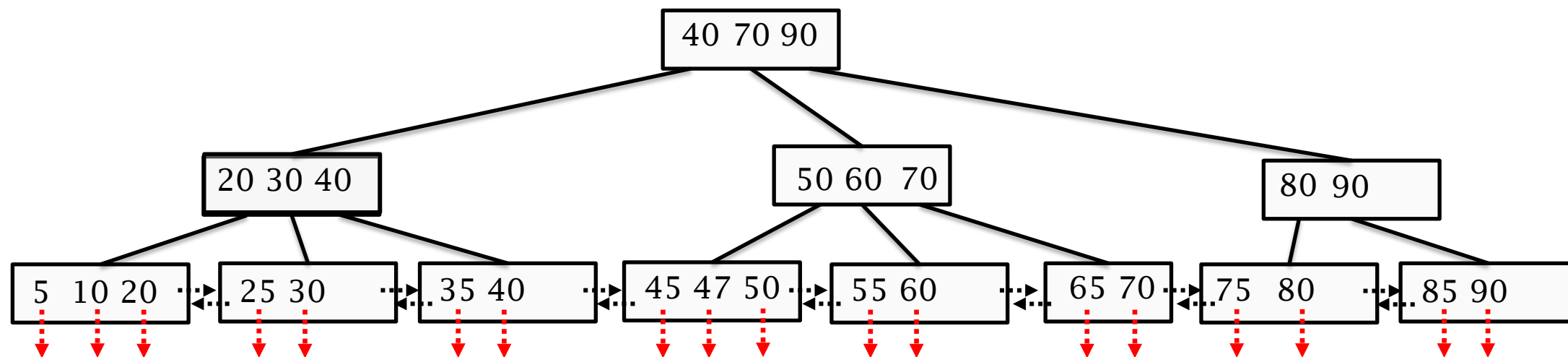
m 阶 B⁺ 树的结构定义如下：

- (1) 每个结点至多有 m 个子结点
- (2) 每个结点(除根外)至少有 $\lceil m / 2 \rceil$ 个子结点
- (3) 根结点至少有两个子结点
- (4) 有 k 个子结点的结点必有 k 个关键码





3 阶 B⁺ 树的例子 (一般阶 ≥ 3)



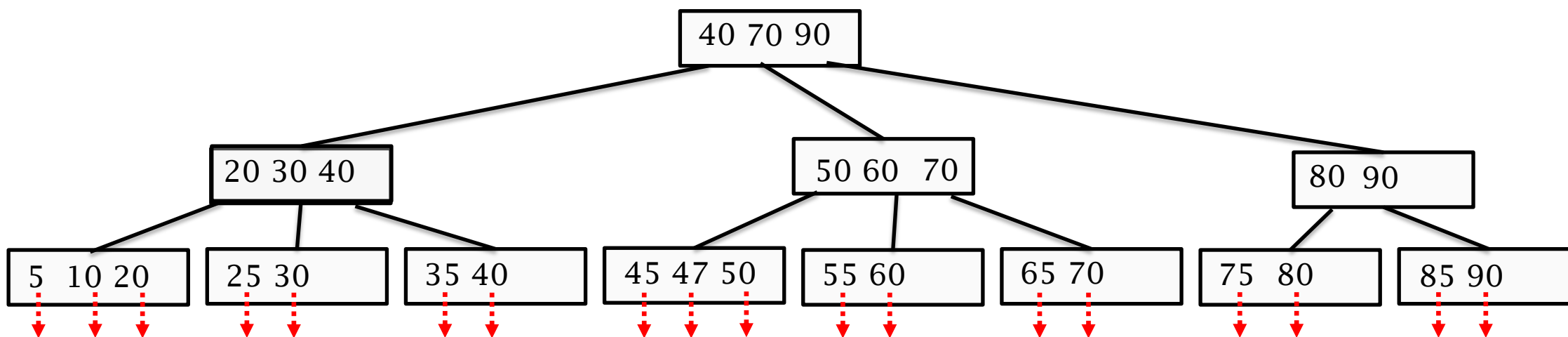
• B⁺ 树的查找

- 在上层已找到待查的关键码，并不停止
- 而是继续沿指针向下一层直到叶结点层的这个关键码

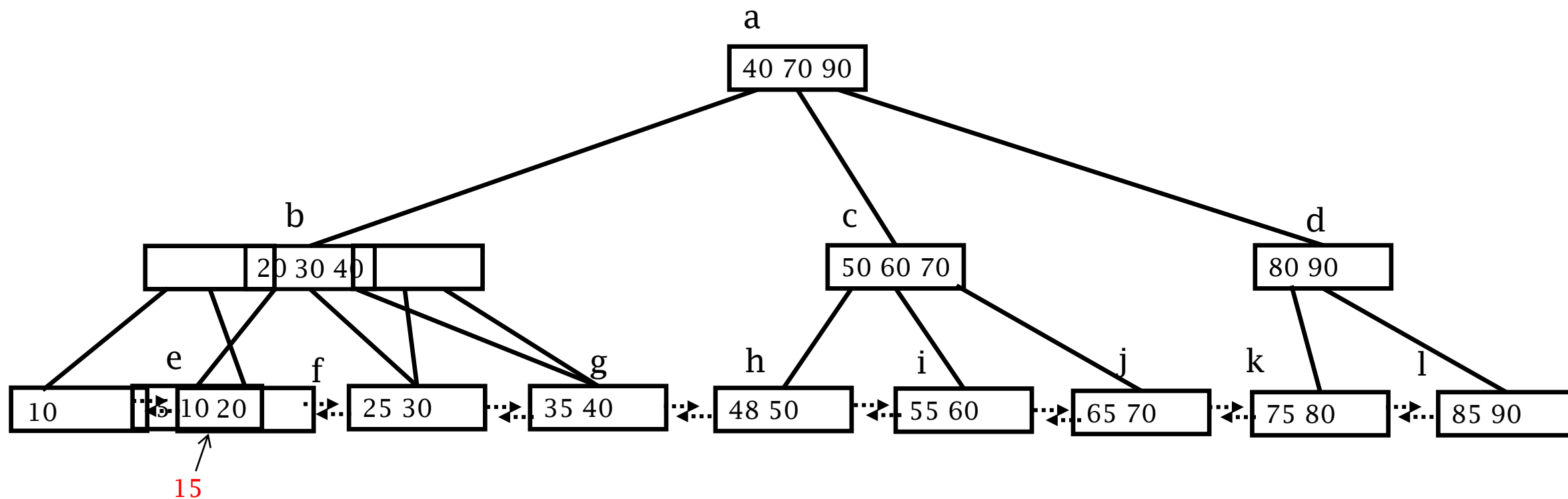


B⁺ 树的插入

- 插入——分裂
 - 过程和 B 树类似
 - 注意保证上一层结点中有这两个结点的最大关键码（或最小关键码）

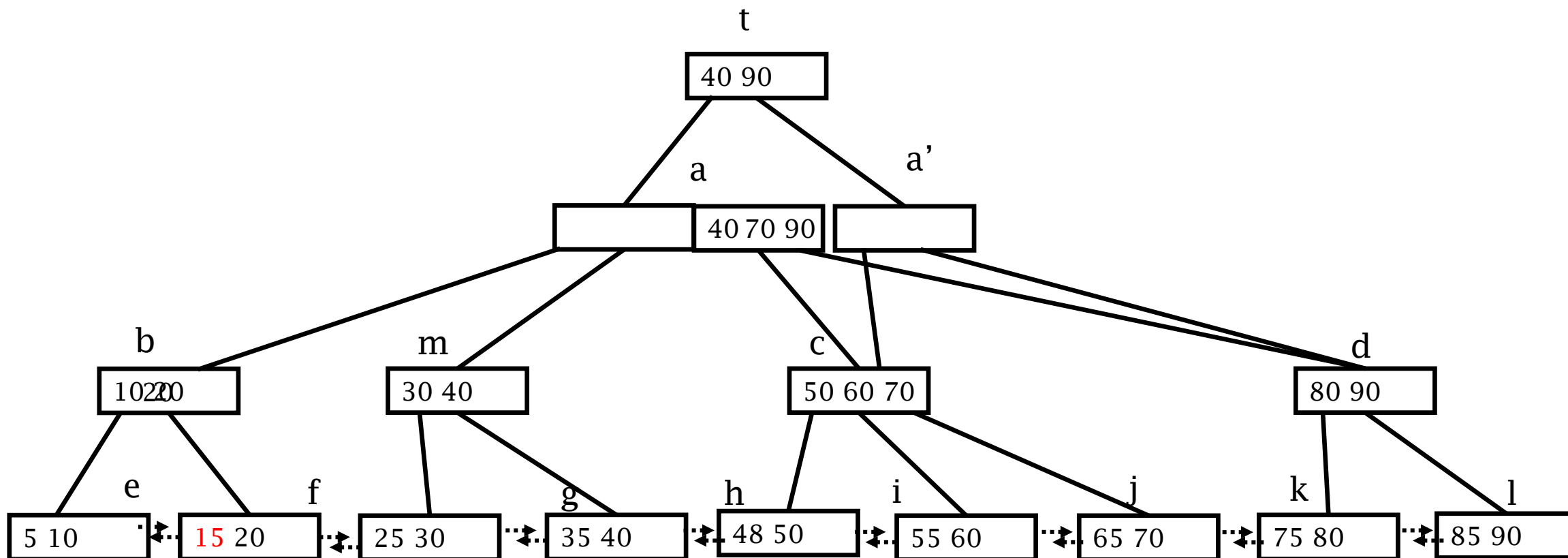


3 阶 B⁺ 树插入15



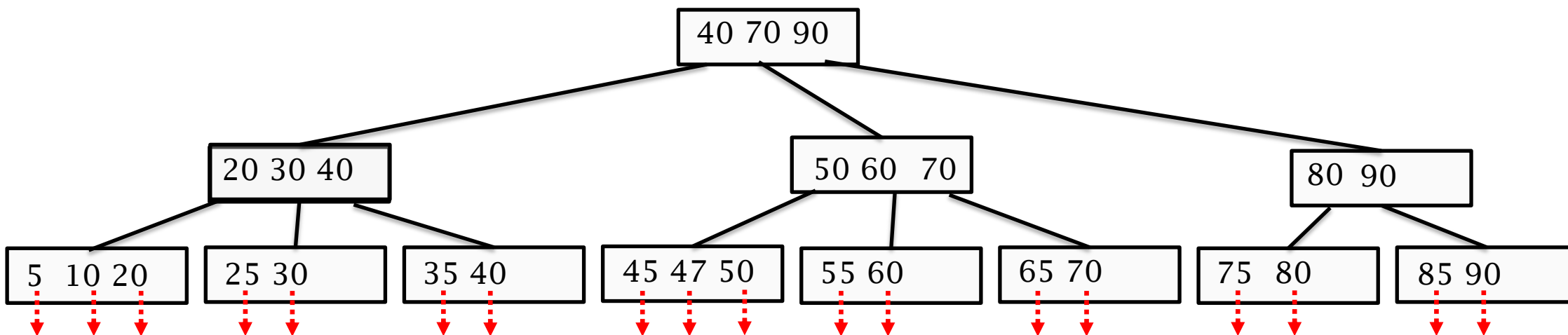
11.4.3 B⁺ 树

在上图 3 阶 B⁺ 树中插入 15 后，树增高一层



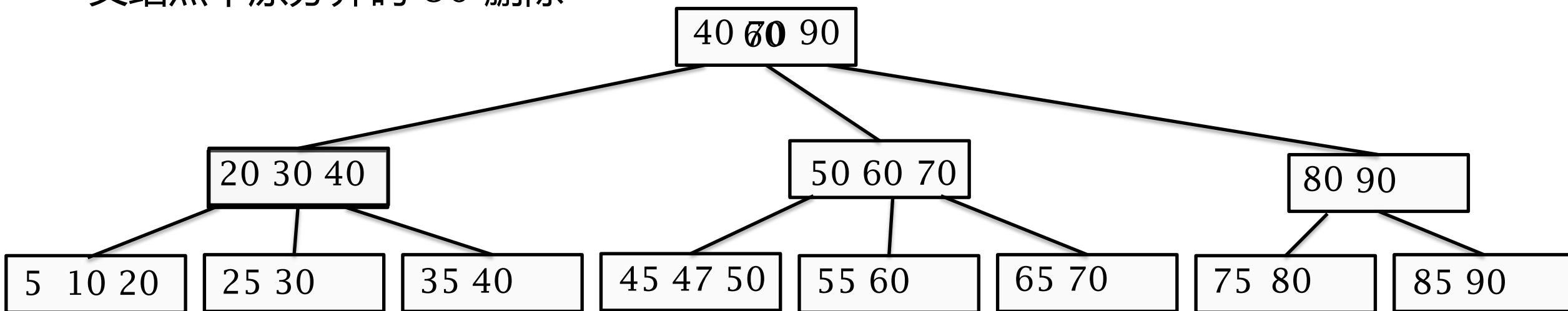
B⁺ 树的删除

- 当关键码下溢出时，与左或右兄弟进行调整（甚至合并）
- 关键码在叶结点层删除后，其在上层的复本可以**保留**，
作为一个“分界关键码”存在
 - 也可以替换为新的最大关键码（或最小关键码）



在 3 阶 B⁺ 树中删除 75

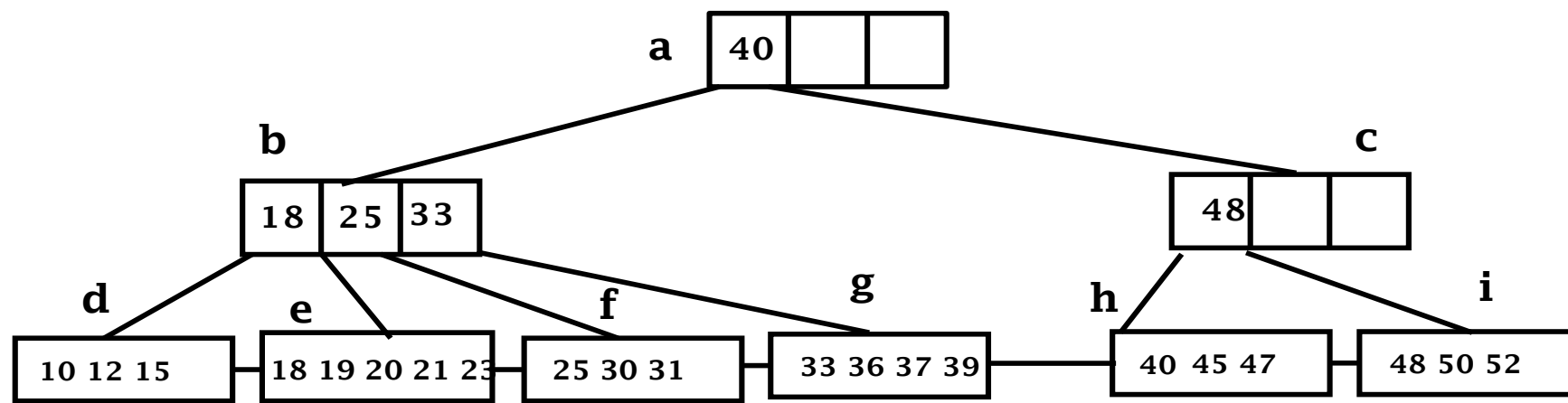
删除 75，发生下溢出，剩余关键码 80 与
 父结点中原分界码 80 修改为 96 (85, 90)
 父结点中原分界码 80 删除





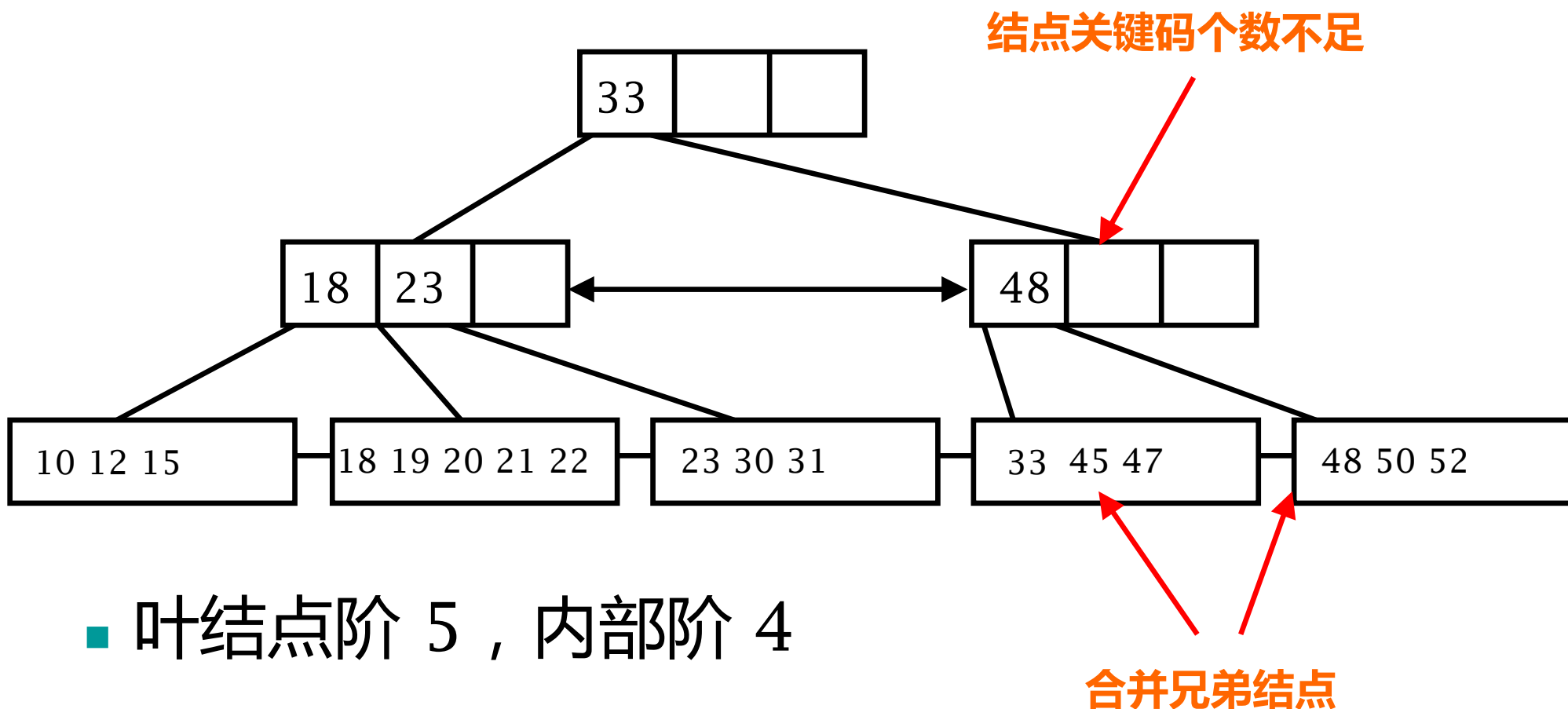
另一种 B⁺ 树

- 叶结点中关键码数目与非叶的不同
 - 内部非叶结点构成 B 树
 - 叶的阶与 B⁺ 树一致
 - 例如，叶结点阶 5，内部阶 4



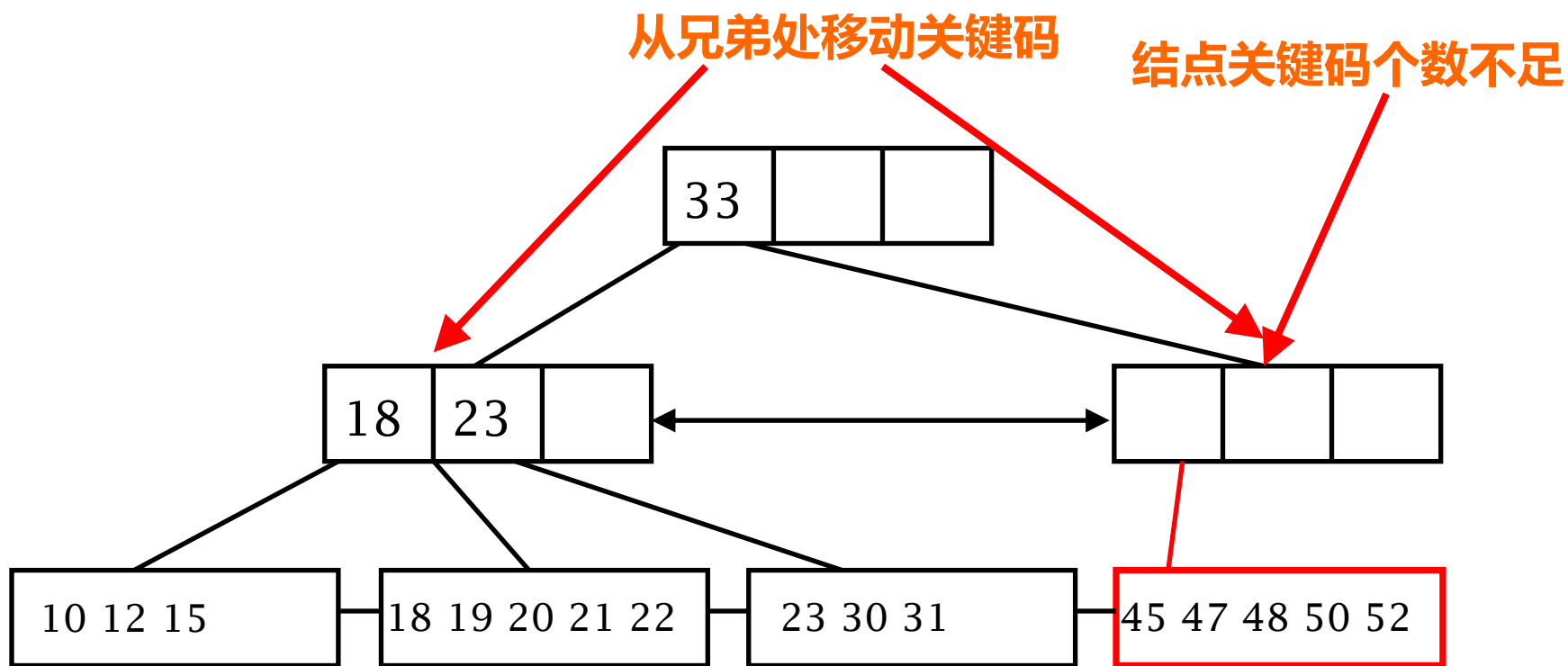


从 B⁺ 树删除关键码值为 33 的记录





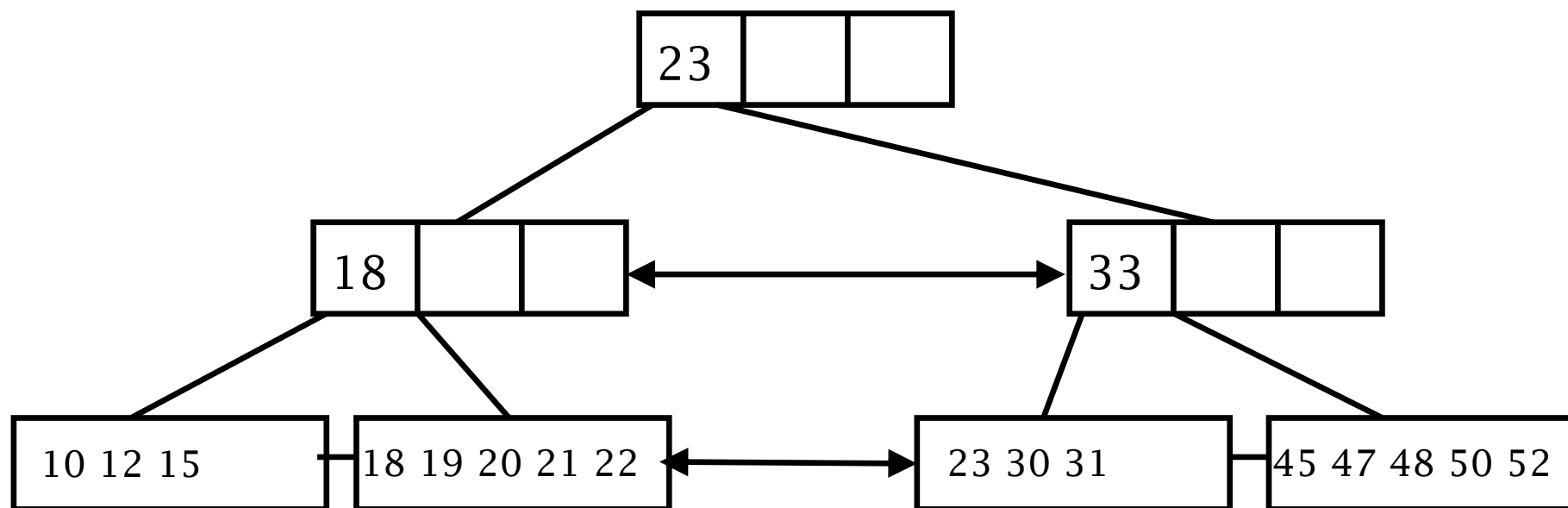
从 B⁺ 树删除关键码值为 33 的记录



■ 叶结点阶 5，内部阶 4



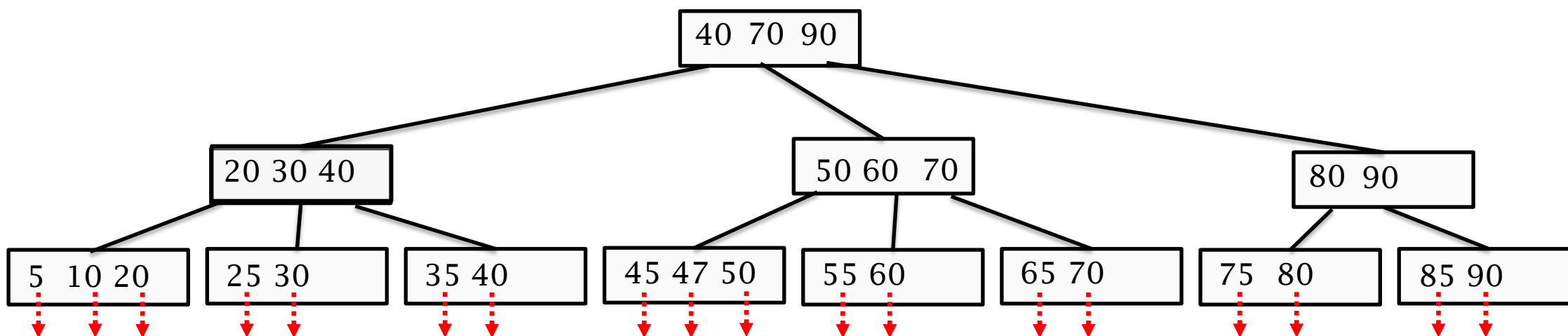
从 B⁺ 树删除关键码值为 33 的记录



■ 叶结点阶 5，内部阶 4

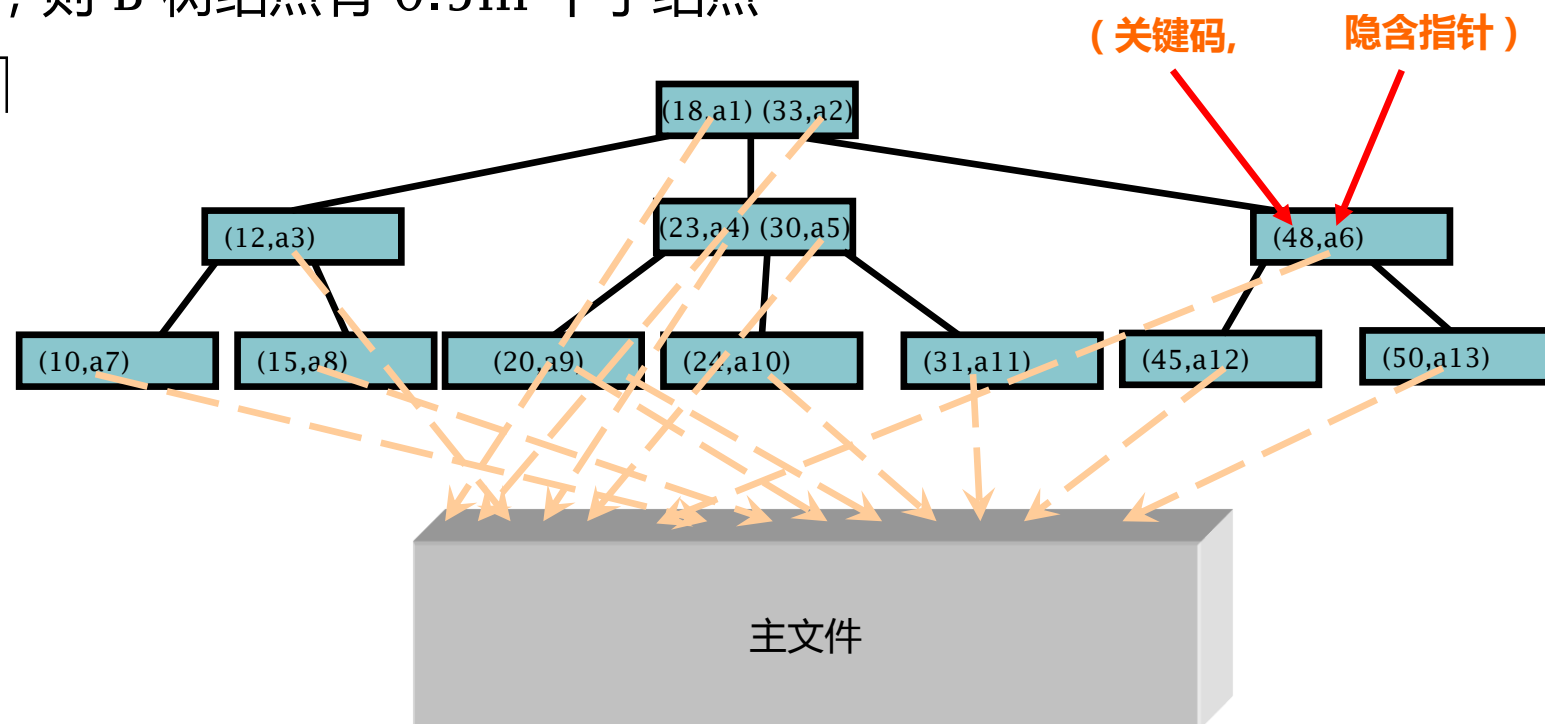
B⁺ 树的存储效率实际上更高

- 假设一个主文件有 N 个记录，假设一个页块可以存 m 个(关键码，子结点页块地址)二元对
- 假设 B⁺ 树平均每个结点有 $0.75m$ 个子结点
 - 充盈度为 $(1+0.5)/2 = 75\%$
- 因此 B⁺ 树的高度为 $\lceil \log_{0.75m} N \rceil$



11.4.4 B 树 B⁺ 树性能比较

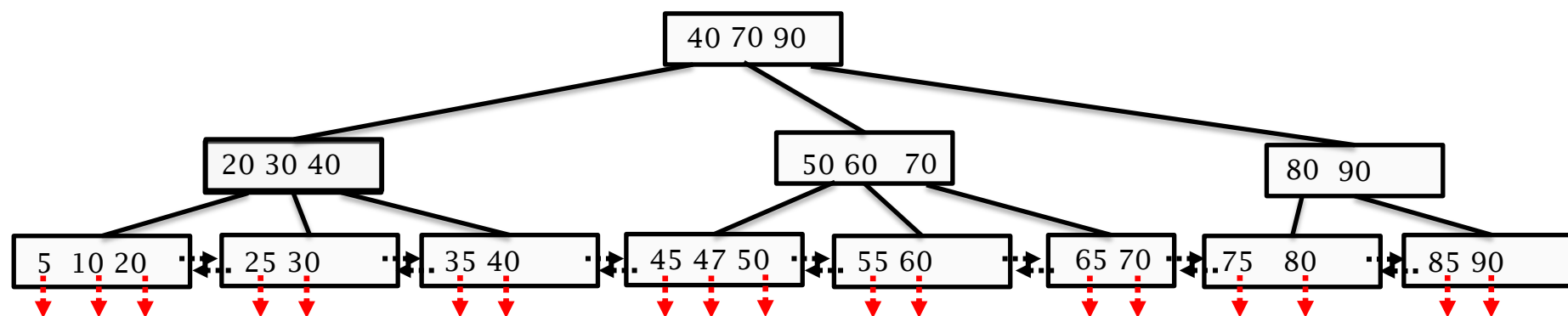
- 可以容纳 m 个(关键码, 子结点页块指针), 假设关键码所占字节数与指针相同
 - 可以容纳 B 树的(关键码, 隐含指针, 子结点页块指针)最多为 $2m/3$ (B 树为 $0.67m$ 阶)。
- 假设 B 树充盈度也是 75%, 则 B 树结点有 $0.5m$ 个子结点
- B 树的高度为 $\lceil \log_{0.5m} N \rceil$



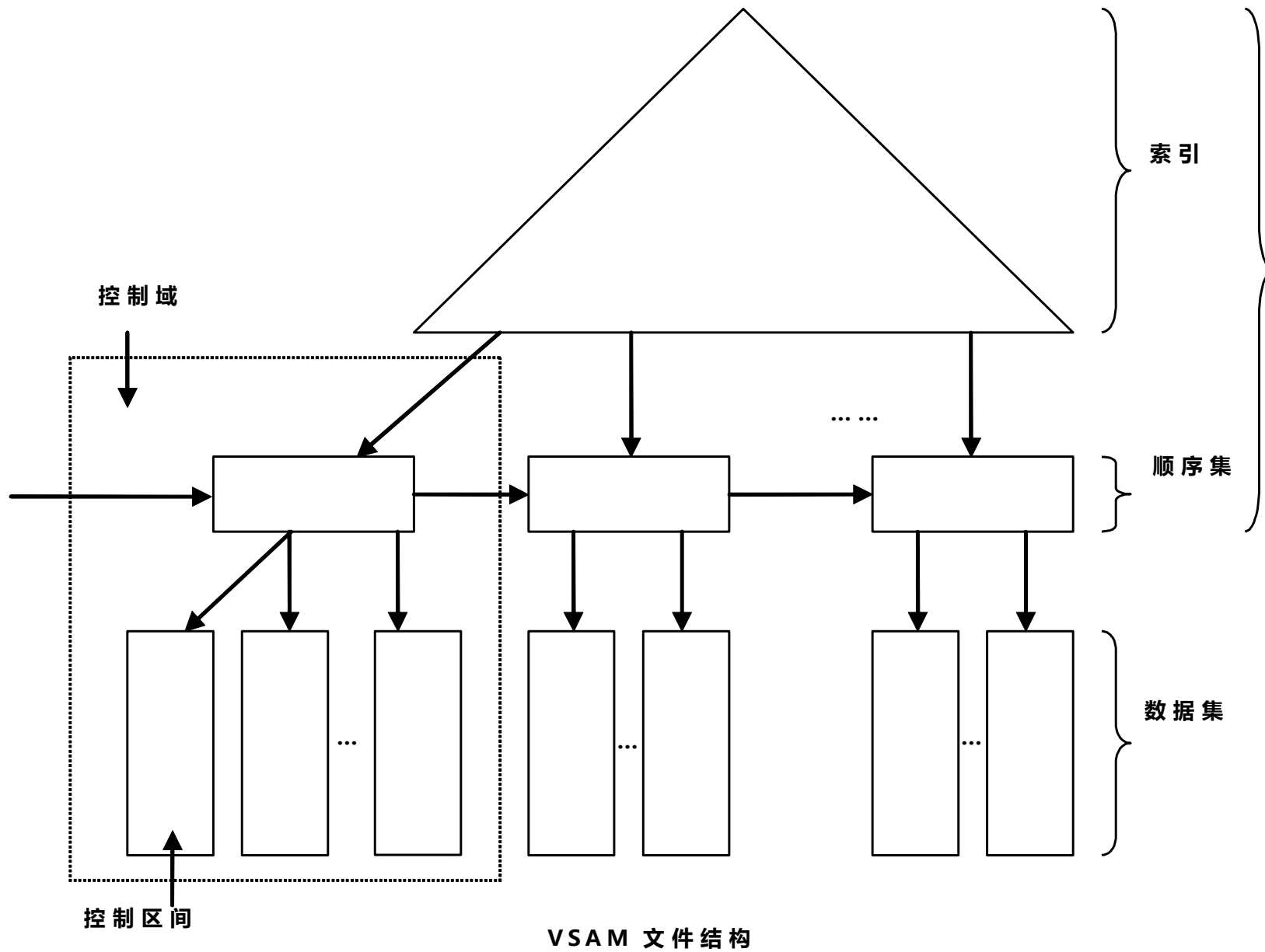


B⁺ 树应用得更为广泛

- B⁺ 树的存储效率更高、检索层次更少（树较矮）
- 因此，B⁺ 树应用得更为广泛
 - 数据库系统主码（primary key）索引
 - 基于B⁺树的磁盘文件虚拟存储存取管理 VSAM (Virtual Storage Access Method)，取代了基于多分树的 ISAM



VSAM 的组成





思考

- 1. 是否存在 2 阶 B^+ 树？
- 2. 为什么相比于 B^+ 树，B 树存储效率低？
- 3. 查阅数据库的相关文献，看看 B^+ 树的作用。



数据结构与算法

谢谢聆听

国家精品课“数据结构与算法”

<http://www.jpk.pku.edu.cn/pkujpk/course/sjjg/>

张铭，王腾蛟，赵海燕

高等教育出版社，2008. 6。“十一五”国家级规划教材