

9. 词典

排解冲突

开放散列

Every mistake I've ever made
Has been rehashed and then replayed
As I got lost along the way.

邓俊辉

deng@tsinghua.edu.cn

多槽位

❖ multiple slots

桶单元细分成若干槽位 `slot`

存放（与同一单元）冲突的词条

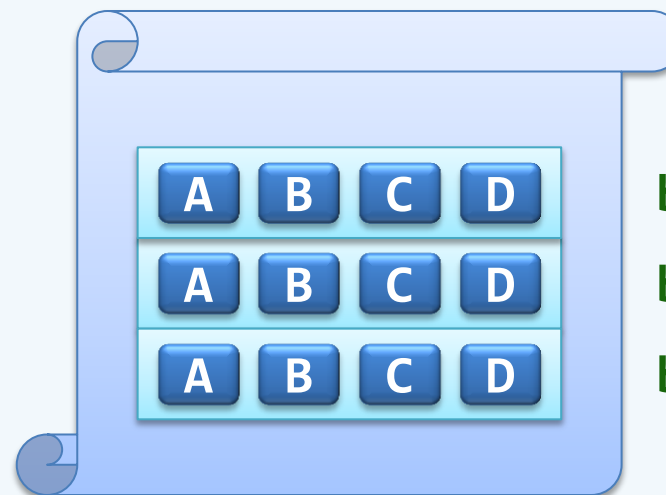
❖ 只要槽位数目不多

依然可以保证 $O(1)$ 的时间效率

❖ 但是，需要为每个桶配备多少个槽，方能保证 $O(1)$ ？ // 难以预测

预留过多，空间 `浪费`

无论预留多少，极端情况下仍有可能 `不够`



`bucket[n + 1]`

`bucket[n]`

`bucket[n - 1]`

独立链

❖ linked-list chaining / separate chaining

每个桶存放一个指针

冲突的词条，组织成列表

❖ 优点 无需为每个桶预备多个槽位

任意多次的冲突都可解决

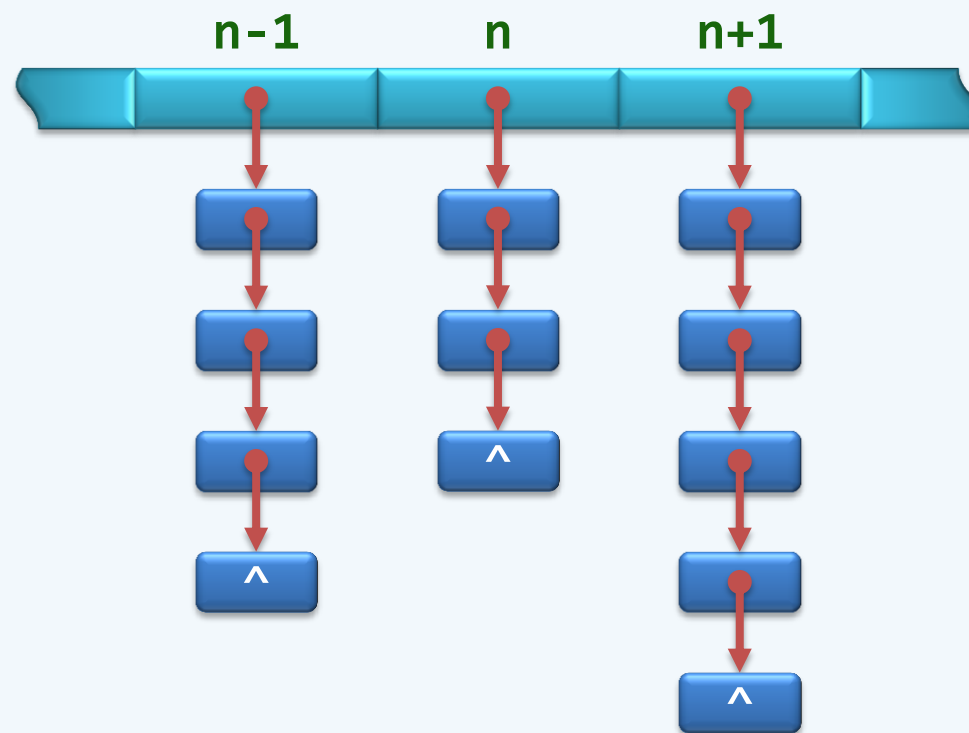
删除操作实现简单、统一

❖ 但是 指针需要额外空间

节点需要动态申请

更重要的是...

❖ 空间未必连续分布，系统缓存几乎失效



公共溢出区

❖ overflow area

单独开辟一块连续空间

发生冲突的词条，顺序存入此区域

❖ 结构简单

算法易于实现

❖ 但是，不冲突则已，一旦发生冲突

最坏情况下，处理冲突词条所需的时间正比于溢出区的规模

