

现有长度为 5，初始为空的散列表  $HT$ ，散列函数  $H(k) = (k+4) \bmod 5$ ，用线性探查再散列法解决冲突。若将关键字序列 2022, 12, 25 依次插入  $HT$  中，然后删除关键字 25，则  $HT$  中查找失败的平均查找长度为

- A. 1                      B. 1.6                      C. 1.8                      D. 2.2

模拟插入和删除流程：

- 插入 2022，映射  $H(2022) = 1$ ，位置空闲，直接插入位置  $h = 1$
- 插入 12，映射  $H(12) = 1$ ，位置被占用，线性探测插入位置  $h = 1 + 1 = 2$
- 插入 25，映射  $H(25) = 4$ ，位置空闲，直接插入位置  $h = 4$
- 删除 25，映射  $H(25) = 4$ ，将此处标记为「已删除」（查找时记为非空闲）

0	1	2	3	4
	2022	12		#

接下来模拟查找流程：设查找数映射的值为  $h$ ，因为是线性探查再散列，所以查找位置序列为： $h \rightarrow (h+1) \rightarrow (h+2) \rightarrow \dots$ 。由于要求「查找失败」的次数，因此必定不会成功（表中的数必定不会等于查找数），依次查找直到遇到「空闲」的位置就停下即可。

- 假如  $h = 0$ ，则查找次数为 1；
- 假如  $h = 1$ ，则查找次数为 3（挪到  $h = 1 \rightarrow 2 \rightarrow 3$  才空闲）；
- 假如  $h = 2$ ，则查找次数为 2（挪到  $h = 2 \rightarrow 3$  才空闲）；
- 假如  $h = 3$ ，则查找次数为 1；
- 假如  $h = 4$ ，则查找次数为 2（挪到  $h = 4 \rightarrow 0$  才空闲）。

对上述 5 种可能的情况作出加权平均即可：

$$ASL_{\text{失败}} = \frac{1+3+2+1+2}{5} = 1.8$$


这个题到此已经做出来了。不过接下来可以简单讲一下为什么删除元素要将格子标记为「已删除」而非直接删除。

比如这回散列表长度为 5， $H(k) = k \bmod 5$ ，线性探查再散列避免冲突。先后插入 3 和 8，得到的散列表肯定是这样的：

0	1	2	3	4
			3	8


现在要删除 3，查找 8，显然 8 在  $h = 4$  位置能查找到。假如把  $h = 3$  标记为「已删除」，那么查找时，先查到  $h = 3$  会视为非空闲而跳到下一个位置，然后查到  $h = 4$  就能返回「查找成功」。

0	1	2	3	4
			#	8



假如把  $h = 3$  位置直接删除，那么查找时，查找 8 的散列位置  $h = 3$  发现是空闲的，就会直接返回「查找失败」，从而产生错误。

0	1	2	3	4
				8



因此，删除散列表的元素时，不能直接删掉，而是应该打标记。

### 【结论】C

【点评】本题考察散列表的用法，具体包括元素的插入、删除、查找方式，线性探查再散列的规则，以及平均查找长度（ASL）的计算等。本题的关键在于删除元素时不能直接删掉，而应该打标记，否则就会算出  $ASL = \frac{1+3+2+1+1}{5} = 1.4$  了。据说在考研真题中算是难度比较大的一道题了。

