Lab04-Hashing

VE281 - Data Structures and Algorithms, Xiaofeng Gao, TA: Qingmin Liu, Autumn 2019

∗ Please upload your assignment to website. Contact webmaster for any questions.

∗ Name:______________ Student ID:______________ Email: ___________________

1. Given a sequence of inputs 192, 42, 142, 56, 39, 319, 14, insert them into a hash table of size 10. Suppose that the hash function is $h(x) = x\%10$. Show the result for the following implementations:

   (a) Hash table using separate chaining. Assume that the insertion is always at the beginning of each linked list.

   (b) Hash table using linear probing.

   (c) Hash table using quadratic probing.

   (d) Hash table using double hashing, with the second hash function as $h_2(x) = (x + 4)\%7$.

Solution.

(a)

[0]
[1]
[2]→ $\boxed{142}$ → $\boxed{42}$ → $\boxed{192}$
[3]
[4]→ $\boxed{14}$
[5]
[6]→ $\boxed{56}$
[7]
[8]
[9]→ $\boxed{319}$ → $\boxed{39}$

(b)

[0]	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
319		192	42	142	14	56			39

(c)

[0]	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
319		192	42	14		142	56		39

(d)

[0]	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
56	319	192		14		42		142	39

□

2. Show the result of rehashing the four hash tables in the Problem 1. Rehash using a new table size of 14, and a new hash function $h(x) = x\%14$. (Hint: The order in rehashing depends on the order stored in the old hash table, not on their initial inserting order.)

Solution.

(a)

[0]→ $\boxed{56}$ → $\boxed{14}$ → $\boxed{42}$
[1]
[2]→ $\boxed{142}$
[3]
[4]
[5]
[6]
[7]
[8]
[9]
[10]→ $\boxed{192}$
[11]→ $\boxed{39}$ → $\boxed{319}$
[12]
[13]

(b)

[0]	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]
42	14	142	56							192	319	39	

(c)

[0]	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]
42	14	142		56						192	319	39	

(d)

[0]	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]
56		142		14				42		192	319	39	

$\square$

3. Suppose we want to design a hash table containing at most 900 elements using linear probing. We require that an unsuccessful search needs no more than 8.5 compares and a successful search needs no more than 3 compares on average. Please determine a proper hash table size.

Solution.
Since for linear probing, we have:

$$U(L) = \frac{1}{2}\left[1 + \left(\frac{1}{1-L}\right)^2\right] \leq 8.5$$

$$S(L) = \frac{1}{2}\left[1 + \frac{1}{1-L}\right] \leq 3$$

Then,

$$L \leq \frac{3}{4} \quad \text{and} \quad L \leq \frac{4}{5},$$

Thus,

$$L = \frac{|S|}{n} \leq \frac{900}{n} \leq \frac{3}{4}.$$

Finally,

$$n \geq \frac{4 \times 900}{3} = 1200.$$

$\square$

4. Implement queues with two stacks. We know that stacks are first in last out (FILO) and queues are first in first out (FIFO). We can implement queues with two stacks. The method is as follows:

   - For **enqueue** operation, push the element into stack S_1.
   - For **dequeue** operation, there are two cases:
     - $S_2 = \emptyset$**,** pop all elements in S_1, push these elements into S_2, pop S_2
     - $S_2 \neq \emptyset$**,** pop S_2

   Using amortized analysis to calculate the complexity of **enqueue** and **dequeue** step.

   Solution.
   Basic idea: all operations have the same amortized cost $\frac{1}{n}\sum_{i=1}^{n}\widehat{C}_i$

   Key observation: $\#pop(S_2) \leq \#push(S_2) \leq \#pop(S_1) \leq \#push(S_1)$

   Thus, we have:

$$
\begin{aligned}
T(n) &= \sum_{i=1}^{n} C_i \\
&= \#pop(S_2) + \#push(S_2) + \#pop(S_1) + \#push(S_1) \\
&\leq 4 \times \#push(S_1) \\
&\leq 4n
\end{aligned}
$$

   Conclusion: on average, each operation takes $O(1)$ time. □