Section 04: Hashing and Math Review

Hashing Problems

1. Hash Table Insertion!

(a) Suppose we have a hash table implemented using separate chaining. This hash table has an internal capacity of 10. Its buckets are implemented using a linked list where new elements are appended to the end. Do not worry about resizing.

Show what this hash table internally looks like after inserting the following key-value pairs in the order given using the hash function $h(x) = 4x$:

$$(1, a), (4, b), (2, c), (17, d), (12, e), (9, e), (19, f), (4, g), (8, c), (12, f)$$

(b) Consider the following scenario:

Suppose we have a hash table with an initial capacity of 12. We resize the hash table by doubling the capacity. Suppose we insert integer keys into this table using the hash function $h(x) = 4x$:

Why is this choice of hash function and initial capacity suboptimal? How can we fix it?

2. More Hash Table Insertion!

For each problem, insert the given elements into the described hash table. Do not worry about resizing the internal array.

(a) Suppose we have a hash table that uses separate chaining and has an internal capacity of 12. Assume that each bucket is a linked list where new elements are added to the front of the list.

Insert the following elements in the EXACT order given using the hash function $h(x) = 4x$:

0, 4, 7, 1, 2, 3, 6, 11, 16

(b) Suppose we have a hash table that uses linear probing and has an internal capacity of 13.

Insert the following elements in the EXACT order given using the hash function $h(x) = 3x$:

2, 4, 6, 7, 15, 13, 19

(c) Suppose we have a hash table that uses quadratic probing and has an internal capacity of 10.

Insert the following elements in the EXACT order given using the hash function $h(x) = x$:

0, 1, 2, 5, 15, 25, 35

(d) Suppose we have a hash table with an initial capacity of 8 using quadratic probing. We resize the hash table by doubling the capacity.

Suppose we insert the integer keys $2^{20}, 2 \cdot 2^{20}, 3 \cdot 2^{20}, 4 \cdot 2^{20}, \ldots$ using the hash function $h(x) = x$.

Describe what the runtime of the dictionary operations will over time as you keep adding these keys to the table.

3. Even More Hash Table Insertion!

(a) Consider the following key-value pairs.

(6, a), (29, b), (41, d). (34, e), (10, f), (64, g), (50, h)

Suppose each key has a hash function $h(k) = 2k$. So, the key 6 would have a hash code of 12. Insert each key-value pair into the following hash tables and draw what their internal state looks like:

(i) A hash table that uses separate chaining. The table has an internal capacity of 10. Assume each bucket is a linked list, where new pairs are appended to the end. Do not worry about resizing.

(ii) A hash table that uses linear probing, with internal capacity 10. Do not worry about resizing.

(iii) A hash table that uses quadratic probing, with internal capacity 10. Do not worry about resizing.

4. Hashing and Mutation

For the following problems, assume that:

1. IntList is a list of integers.
2. The hash code of an IntList is the sum of the integers in the list.
3. IntLists are considered equal only if they have the same size and the same values in the same order.
4. FourBucketHashMap uses separate chaining and the new items are added to the back of each bucket.
5. FourBucketHashMap always has four buckets and never resizes.

Consider the following code:

```
FourBucketHashMap<IntList, String> fbhm = new FourBucketHashMap<>();
IntList list1 = IntList.of(1, 2);
fbhm.put(list1, "dog");
// Part i
list1.add(3);
// Part ii
```

(a) At Part i (line 4), what will be returned from the following statement?

```
fbhm.get(IntList.of(1, 2));
```

(b) At Part II (line 6), what will be returned from the following statements?

```
fbhm.get(IntList.of(1, 2));
fbhm.get(IntList.of(1, 2, 3));
```

(c) Is there a problem with the code? If so, explain.

5. Debugging a Hash Table

Suppose we are in the process of implementing a hash map that uses open addressing and quadratic probing and want to implement the `delete` method.

(a) Consider the following implementation of `delete`. List every bug you can find.

Note: You can assume that the given code compiles. Focus on finding run-time bugs, not compile-time bugs.

```
1   public class QuadraticProbingHashTable<K, V> {
2       private Pair<K, V>[] array;
3       private int size;
4
5       private static class Pair<K, V> {
6           public K key;
7           public V value;
8       }
9
10      // Other methods are omitted, but functional.
11
12      /**
13       * Deletes the key-value pair associated with the key, and
14       * returns the old value.
15       *
16       * @throws NoSuchKeyException if the key-value pair does not exist in the method.
17       */
18      public V delete(K key) {
19          int index = key.hashCode() % this.array.length;
20
21          int i = 0;
22          while (this.array[index] != null && !this.array[index].key.equals(key)) {
23              i += 1;
24              index = (index + i * i) % this.array.length;
25          }
26
27          if (this.array[index] == null) {
28              throw new NoSuchKeyException("Key-value pair not in dictionary");
29          }
30
31          this.array[index] = null;
32
33          return this.array[index].value;
34      }
35  }
```

(b) Let's suppose the `Pair` array has the following elements (pretend the array fit on one line):

["lily", V_2]	["castle", V_6]	["resource", V_1]	["hard", V_9]	["bathtub", V_0]
["wage", V_4]	["refund", V_7]	["satisfied", V_6]	["spring", V_8]	["spill", V_3]

And, that the following keys have the following hash codes:

Key	Hash Code
“bathtub”	9744
“resource”	4452
“lily”	7410
“spill”	2269
“wage”	8714
“castle”	2900
“satisfied”	9251
“refund”	8105
“spring”	6494
“hard”	9821

What happens when we call `delete` with the following inputs? Be sure write out the resultant array, and to do these method calls *in order*. (**Note**: If a call results in an infinite loop or an error, explain what happened, but don't change the array contents for the next question.)

(i) `delete(“lily”)`

(ii) `delete(“spring”)`

(iii) `delete(“castle”)`

(iv) `delete(“bananas”)`

(v) `delete(null)`

(c) List four different test cases you would write to test this method. For each test case, be sure to either describe or draw out what the table's internal fields look like, as well as the expected outcome (assuming the `delete` method was implemented correctly). **Hint**: You may use the inputs previously given to help you identify tests, but it's up to you to describe what kind of input they are testing generally.

Math Review

6. Tree Method

Find a summation for the total work of the following expressions using the Tree Method.

Hint: Just as a reminder, here are the steps you should go through for **any** Tree Method Problem:

i. Draw the recurrence tree.

ii. What is the size of the **input** to each node at level i? As in class, we call the root level $i = 0$. This means that at $i = 0$, your expression for the input should equal n.

iii. What is the amount of **work** done by each node at the i-th *recursive* level?

iv. What is the total number of nodes at level i? As in class, we call the root level $i = 0$. This means that at $i = 0$, your expression for the total number of nodes should equal 1.

v. What is the total work done across the i-th *recursive* level?

vi. What value of i does the last level of the tree occur at?

vii. What is the total work done across the base case level of the tree (i.e. the last level)?

viii. Combine your answers from previous parts to get an expression for the total work.

(a) $T(n) = \begin{cases} 2T(n/3) + 5n & \text{if } n > 1 \\ 9 & \text{otherwise} \end{cases}$

(b) $T(n) = \begin{cases} T(n-1) + n^2 & \text{if } n > 19 \\ 57 & \text{otherwise} \end{cases}$

(c) $T(n) = \begin{cases} T(n/2) + n^2 & \text{if } n \geq 4 \\ 5 & \text{otherwise} \end{cases}$

7. Best case and worst case runtimes

For the following code snippet give the big-Θ bound on the worst case runtime as well the big-Θ bound on the best case runtime, in terms of n the size of the input array.

```
1  void print(int[] input) {
2      int i = 0;
3      while (i < input.length - 1) {
4          if (input[i] > input[i + 1]) {
5              for (int j = 0; j < input.length; j++) {
6                  System.out.println("uh I don't think this is sorted plz help");
7              }
8          } else {
9              System.out.println("input[i] <= input[i + 1] is true");
10         }
11         i++;
12     }
13 }
```

8. Big-O, Big-Omega True/False Statements

For each of the statements determine if the statement is true or false. You do not need to justify your answer.

(a) $n^3 + 30n^2 + 300n$ is $\mathcal{O}(n^3)$

(b) $nlog(n)$ is $\mathcal{O}(log(n))$

(c) $n^3 - 3n + 3n^2$ is $\mathcal{O}(n^2)$

(d) 1 is $\Omega(n)$

(e) $.5n^3$ is $\Omega(n^3)$

9. Eyeballing Big-Θ bounds

For each of the following code blocks, what is the worst-case runtime? Give a big-Θ bound. You do not need to justify your answer.

(a)
```
void f1(int n) {
    int i = 1;
    int j;
    while(i < n*n*n*n) {
        j = n;
        while (j > 1) {
            j -= 1;
        }
        i += n;
    }
}
```

(b)
```
int f2(int n) {
    for (int i = 0; i < n; i++) {
        for (int j = 0; j < n; j++) {
            System.out.println("j = " + j);
        }
        for (int k = 0; k < i; k++) {
            System.out.println("k = " + k);
            for (int m = 0; m < 100000; m++) {
                System.out.println("m = " + m);
            }
        }
    }
}
```

(c)
```
int f3(n) {
    count = 0;
    if (n < 1000) {
        for (int i = 0; i < n; i++) {
            for (int j = 0; j < n; j++) {
                for (int k = 0; k < i; k++) {
                    count++;
                }
            }
        }
    } else {
        for (int i = 0; i < n; i++) {
            count++;
        }
    }
    return count;
}
```

(d)
```
void f4(int n) {
    // NOTE: This is your data structure from the first project.
    LinkedDeque<Integer> deque = new LinkedDeque<>();
    for (int i = 0; i < n; i++) {
        if (deque.size() > 20) {
            deque.removeFirst();
        }
        deque.addLast(i);
    }
    for (int i = 0; i < deque.size(); i++) {
        System.out.println(deque.get(i));
    }
}
```

10. Modeling

Consider the following method. Let n be the integer value of the `n` parameter, and let m be the size of the `LinkedDeque`.

```
public int mystery(int n, LinkedDeque<Integer> deque) {
    if (n < 7) {
        System.out.println("???");
        int out = 0;
        for (int i = 0; i < n; i++) {
            out += i;
        }
        return out;
    } else {
        System.out.println("???");
        System.out.println("???");
        out = 0;
        // NOTE: Assume LinkedDeque has working, efficient iterator.
        for (int i : deque) {
            out += 1;
            for (int j = 0; j < deque.size(); j++) {
                System.out.println(deque.get(j));
            }
        }
        return out + 2 * mystery(n - 4, deque) + 3 * mystery(n / 2, deque);
    }
}
```

Give a recurrence formula for the **worst-case** running time of this code. It's OK to provide a $\mathcal{O}$ for non-recursive terms (for example if the running time is $A(n) = 4A(n/3) + 25n$, you need to get the 4 and the 3 right but you don't have to worry about getting the 25 right). Just show us how you got there.