

Lecture 09: Hashing Part 1 (Separate Chaining)

1. Assume Point class has implemented both `equals()` and `hashCode()` correctly. Draw the pointSet and depict what is happening with above code.

```
HashSet<Point> pointSet = new HashSet<>();
Point p1 = new Point(3,5); pointSet.add(p1);

System.out.println(pointSet.contains(p1)); //outputs?

Point p2 = new Point(3,5);
System.out.println(pointSet.contains(p2)); //outputs?

p2.translate(2,1);
System.out.println(pointSet.contains(new Point(5,6))); //outputs?

System.out.println(pointSet.contains(p2)); //outputs?

System.out.println(pointSet.contains(new Point(3,5))); //outputs?
```

Hash Table

1. Hash table's ADT

-
-
-

2. What is ADT? It defines a set of _____ it supports (and values it can have).

ADT defines _____ a data structure does (behavior), while the actual data structure defines _____ it does it (implementation).

3. Notation

U: _____ of ALL possible keys

K: Set of keys actually _____ in the dictionary

$|K| = n$: the number of current keys stored m: the capacity of the hash table

4. We should do hashing only when _____. Why?
Table size m is typically proportional to _____.

5. Hash function maps U to an index $0 \dots m-1$.
Hash function must be deterministic, fast, and has great uniformity.

CS314H CSB FA15 Class Notes
Lecture 09: Hashing Part 1 (Separate Chaining)

6. Hash determines the _____ at which the key is stored.

Collision resolution: Separate chaining

7. Separate chaining uses _____ to store all collided entries at a given slot.

8.

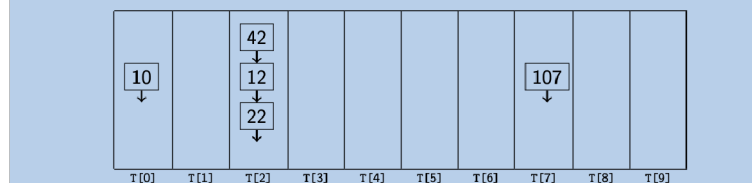
Example (Insert: 10, 22, 107, 12, 42)



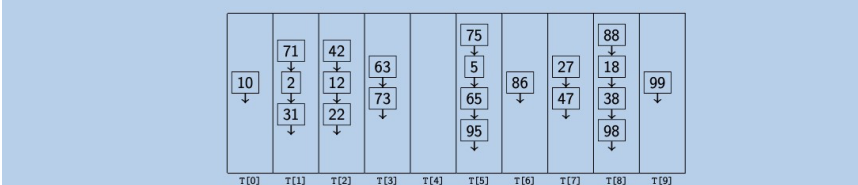
- What is the hash function used here? $h(x) = \underline{\hspace{2cm}}$.
- What is the worst case time complexity of insert?
- What is the worst case time complexity of find?
- Average case?

9. _____ tells “how full” the hash table is. Defined as _____. The number of current elements stored divided by the size (capacity) of the hash table.

Example (What is the Load Factor?)



Example (What is the Load Factor?)



Load factor also tells the _____ number of _____ elements in each slot.

10. Why average case of insert, find, delete is all $O(1)$? Explain in your own word using load factor.

11. ChainingHashSet Implementation.

Why private static class for HashNode?

Why `key.hashCode() % table.length` for computing the hashed index?

Why rehashing?

12. How to improve the performance further?

13. What are the pros/cons of separate chaining?