App Academy

Hashing, Maps, Sets

Hash Sets / Maps

- Hash sets and maps are probably the most common data structures you're ever going to use, whether we're talking about interviews or practical use
- Set - a set of unique values
    - Methods
        - new_set = set()
        - len(new_set)
        - new_set.add()
        - new_set.remove(), new_set.discard()
        - value in new_set
    - Unlike an array, does not have indices
    - Useful when we need to keep track of unique pieces of data that should only appear once
        - e.g. coordinates on a matrix
    - Example: {5,7,9,3,30}

Hash Sets / Maps (cont.)

- Maps and Objects
  - store data in key-value pairs
    - Python dictionary methods
      - dictionary["key"], dictionary.get("key", default_value)
      - dictionary.keys()
      - dictionary.values()
      - dictionary.items()
      - del dictionary["key"]
    - Python Default Dictionary
      - from collections import defaultdict
      - defaultdict_demo = defaultdict(int)
      - Other values for parameter
        - set
        - list

Runtime Complexities

Maps, Objects, and Sets

Operations	Big-O Time
Insert value	O(1)
Remove value	O(1)
Search value	O(1)

As we can see, these hash data structures are extremely efficient, which is why it's so commonly used.

Hash Map Implementation

HashMap	
Index	Key, Value

- Use a hashing function to convert key into an integer, then use that integer as the index
- Minimize collisions
  - We cannot actually completely get rid of collisions! However we can make it extremely unlikely using different strategies
  - rehashing
    - resize array when half full (similar to dynamic arrays)
      - recompute the hash
        - may need to move elements to new indices
  - chaining
    - store linked lists of pairs instead of just key-value pairs
  - open addressing
    - try the next open position
  -
- Demo
  - hashmap["Alice"] = "NYC";
  - hashmap["Brad"] = "Chicago";
  - hashmap["Collin"] = "Seattle";

Questions?

Demos

- Two Sum
- Matrix Set Zeroes
- Group Anagrams

Let’s practice!

- Review:
    - Valid Sudoku
    - Longest Consecutive Sequence
- Bonus:
    - Design HashSet
    - Design HashMap
    - Design Twitter
    - LRU Cache