

01_Sequences

Arrays and Strings.

<https://www.techinterviewhandbook.org/algorithms/array/>

- [Array Cheatsheet](#)
- [01_Arrays](#)
 - [Brute Force](#)
 - [Two Pointers: Inward Traversal](#)
 - [Two Pointers: Unidirectional Traversal](#)
 - [Two Pointers: Stage Traversal](#)
 - [Sliding Window: Fixed](#)
 - [Sliding Window: Dynamic](#)
 - [Traversing Array From The Right](#)
 - [Sorting The Array](#)
 - [Index as a Hash Key](#)
- [String Cheatsheet](#)
- [02_Strings](#)
 - [Counting Characters in a String](#)
 - [String of Unique Characters](#)
 - [Anagram](#)
 - [Palindrome](#)

Array Cheatsheet

Arrays

Arrays hold values of the same type at contiguous memory locations. In an array, we're usually concerned about two things - the position/index of an element and the element itself.

Advantages

- Store multiple elements of the same type with one single variable name
- Accessing elements is fast ($O(1)$) as long as you have the index, as opposed to linked lists where you have to traverse from the head ($O(n)$).

Disadvantages

- Addition and removal of elements in the middle of an array is slow ($O(n)$) because the remaining elements need to be shifted to accommodate the new/missing element.
 - An exception to this is if the position to be inserted/removed is at the end of the array ($O(1)$).
- It cannot alter its size after initialization.
 - If an insertion causes the total number of elements to exceed the size, a new array has to be allocated and the existing elements have to be copied over.
 - The act of creating a new array and transferring elements over takes $O(n)$ time.

Common terms

- Subarray - A range of contiguous values within an array.
 - Example: given an array `[2, 3, 6, 1, 5, 4]`, `[3, 6, 1]` is a subarray while `[3, 1, 5]` is not a subarray.
- Subsequence - A sequence that can be derived from the given sequence by deleting some or no elements without changing the order of the remaining elements.
 - Example: given an array `[2, 3, 6, 1, 5, 4]`, `[3, 1, 5]` is a subsequence but `[3, 5, 1]` is not a subsequence.

Things to look out for during interviews

- Clarify if there are duplicate values in the array. Would the presence of duplicate values affect the answer? Does it make the question simpler or harder?
- When using an index to iterate through array elements, be careful not to go out of bounds.

- Be mindful about slicing or concatenating arrays in your code. Typically, slicing and concatenating arrays would take $O(n)$ time. Use start and end indices to demarcate a subarray/range where possible.

Corner Cases

- Empty sequence
- Sequence with 1 or 2 elements
- Sequence with repeated elements
- Duplicated values in the sequence

Time Complexity

| Operation | Big-O | Note |
|-----------------------|--------------|--|
| Access | $O(1)$ | |
| Search | $O(n)$ | |
| Search (sorted array) | $O(\log(n))$ | |
| Insert | $O(n)$ | Insertion would require shifting all the subsequent elements to the right by one and that takes $O(n)$ |
| Insert (at the end) | $O(1)$ | Special case of insertion where no other element needs to be shifted |
| Remove | $O(n)$ | Removal would require shifting all the subsequent elements to the left by one and that takes $O(n)$ |
| Remove (at the end) | $O(1)$ | Special case of removal where no other element needs to be |

Resource

[Array in Data Structure: What is, Arrays Operations \[Examples\]](#)

01_Arrays

Brute Force

The brute force approach tries every possible combination to check for a solution, without leveraging any special properties or optimizations of the data (such as sorted order). Typically involves:

- Nested loops
 - Outer Loop traverses the array for the first element of the pair
 - Inner Loop traverses the rest of the array to find second element
- Recursion

Drawbacks:

- **Time Complexity:** Typically $O(n^2)$ or slower, impractical for large datasets.
- **Redundancy:** Many computations are repeated unnecessarily



Target: 35. Starting brute force search...

```
def brute_force_two_sum(nums, target):
    # Iterate through each element in the list
    for i in range(len(nums)):
        # For each element, check every other element that comes after it
        for j in range(i + 1, len(nums)):
            # Check if the current pair sums to the target
            if nums[i] + nums[j] == target:
                return (i, j) # Return the indices as a tuple
    # If no pair is found that adds up to the target, return None
    return None

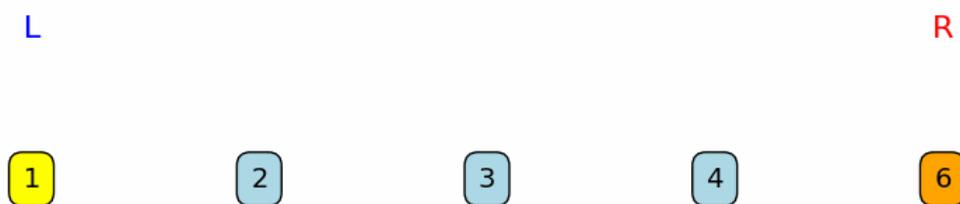
# Example usage:
```

```
nums = [2, 7, 11, 15]
target = 9

result = brute_force_two_sum(nums, target)
if result:
    print("Pair found at indices:", result)
else:
    print("No pair found that adds up to the target.")
```

Two Pointers: Inward Traversal

A pointer is a variable that represents an index or position within a data structure, such as an Array or Linked List. With Two pointers we can make comparisons, with a pointer at two different positions, and infer a decision based on that.



Target: 6. Start: L at index 0 (1), R at index 4 (6)

When to use:

- Data Structure: Linear such as Array, Linked List
- Sorted Array or Palindrome
- Result asks for Pair of Values
- One Result is decided by a Pair of Values

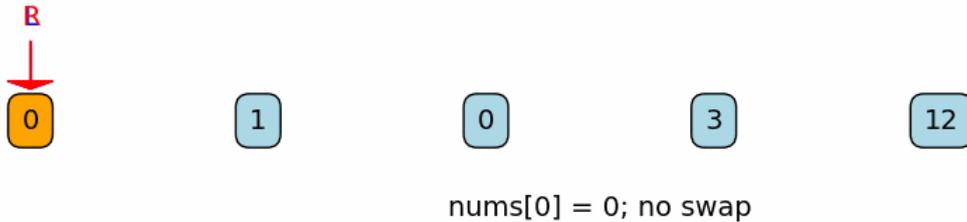
Real-World Example:

- Garbage Collection Algorithm

```
def pair_sum_sorted(nums: List[int], target: int) -> List[int]:
    left, right = 0, len(nums)-1
    while left < right:
        sums = nums[left] + nums[right]
        # if the sum is smaller, increment the left pointer, aiming
        # to increase the sum towards the target value
        if sums < target:
```

```
    left += 1
# if the sum is larger, decrement the right pointer, aiming
# to decrease the sum towards the target value
elif sums > target:
    right -= 1
# If target pair is found return its indices
else:
    return [left, right]
return []
```

Two Pointers: Unidirectional Traversal



```
def shift_zeros_to_the_end(nums: List[int]) -> None:
    # The 'left' pointer is used to position non-zero elements.
    left = 0
    # Iterate through the array using a 'right' pointer to locate non-zero
    # elements.
    for right in range(len(nums)):
        if nums[right] != 0:
            nums[left], nums[right] = nums[right], nums[left]
            # Increment 'left' since it now points to a position already occupied
            # by a non-zero element.
            left += 1
```

Two Pointers: Stage Traversal

L

3

8

5

12

7

4

6

R

Start partitioning by parity: even numbers (L) and odd numbers (R)

Problem: Partition Array by Parity

Description:

Given an integer array `nums`, rearrange the array **in-place** such that all even numbers appear before all odd numbers. The order of the elements within the even or odd group does not matter.

Return the array after rearrangement.

You must solve the problem without using extra space (i.e. in $O(1)$ extra space) and in one pass if possible.

Example 1:

Input: `nums = [3, 8, 5, 12, 7, 4, 6]` Output: `[8, 12, 4, 6, 3, 5, 7]` Explanation: The even numbers `[8, 12, 4, 6]` appear before the odd numbers `[3, 5, 7]`. Note that the order within each group does not matter.

Example 2:

Input: `nums = [1, 3, 5, 7]` Output: `[1, 3, 5, 7]` Explanation: Since there are no even numbers, the array remains unchanged.

Constraints:

- $1 \leq \text{nums.length} \leq 10^5$
- $0 \leq \text{nums}[i] \leq 10^5$

1. Initialization:

The script starts with two pointers:

- `left` at the beginning (index 0)
- `right` at the end (last index)

2. Stage Traversal:

In each loop iteration:

- It prints the current array and the positions (and values) of the two pointers.
- If the number at the left pointer is even, that element is already in the correct half, so the left pointer is moved right.
- If the number at the right pointer is odd, that element is in the correct half, so the right pointer is moved left.
- If neither of those conditions holds (meaning the left number is odd and the right number is even), the two values are swapped. This moves the even number toward the left and the odd number toward the right.

3. Termination:

The loop ends when the left pointer is no longer less than the right pointer. At that point, the array is partitioned with evens on the left and odds on the right.

```
def partition_by_parity(nums):
    left = 0          # Initialize the left pointer at the beginning.
    right = len(nums) - 1  # Initialize the right pointer at the end.
    print("Initial array:", nums)

    while left < right:
        print("\nCurrent array:", nums)
        print(f"Left pointer at index {left} with value {nums[left]}")
        print(f"Right pointer at index {right} with value {nums[right]}")

        # If the left element is even, it's already in the correct partition.
        if nums[left] % 2 == 0:
            print(f"{nums[left]} is even, so move the left pointer right.")
            left += 1

        # If the right element is odd, it's already in the correct partition.
        elif nums[right] % 2 == 1:
            print(f"{nums[right]} is odd, so move the right pointer left.")
            right -= 1

        # Otherwise, the left element is odd and the right element is even.
        # In that case, swap them.
```

```
else:
```

```
    print(f"Swapping {nums[left]} (odd) and {nums[right]} (even).")
```

```
    nums[left], nums[right] = nums[right], nums[left]
```

```
    left += 1
```

```
    right -= 1
```

```
print("\nFinal partitioned array:", nums)
```

```
# Example usage:
```

```
nums = [3, 8, 5, 12, 7, 4, 6]
```

```
partition_by_parity(nums)
```

Sliding Window: Fixed

A subset of the Two Pointer Method, but uses left and right pointers to define the bounds of a "window" in iterable data structures like arrays. The window defines the subcomponent, like subarray or substring, and it slides across the data structure in one direction, searching for a subcomponent that meets a certain requirement.

When to use:

- Data Structure: Linear such as Array, Linked List
- Find a Subcomponent of a length

Brute Force:

- Finding all possible subcomponents for an answer using a Nested Loop
 - Outer Loop traverses the array for the first element of the pair
 - Inner Loop traverses the rest of the array to find second element
- Time Complexity is $O(n^2)$ where n is length of the loop (Two Loops)

Real-World Example:

Buffering in Video Streaming

Sliding Window: Dynamic

Traversing Array From The Right



Target: 4. Start right-to-left traversal.

Find Last Occurrence of Target in Array

Description:

Given an integer array `nums` and an integer `target`, return the index of the last occurrence of `target` in `nums`. If the target is not found, return -1.

You must solve this problem with an efficient $O(n)$ time solution by traversing the array from right to left.

Example 1:

Input: `nums = [1, 3, 5, 3, 2]`, `target = 3`

Output: 3

Explanation: The target 3 appears at indices 1 and 3, but its last occurrence is at index 3.

Example 2:

Input: `nums = [1, 2, 3, 4]`, `target = 5`

Output: -1

Explanation: The target 5 is not present in the array.

Constraints:

- $1 \leq \text{nums.length} \leq 10^5$
- $-10^9 \leq \text{nums}[i] \leq 10^9$

- $-10^9 \leq \text{target} \leq 10^9$

```
from typing import List

def find_last_occurrence(nums: List[int], target: int) -> int:
    """
    Returns the index of the last occurrence of target in nums.
    If the target is not found, returns -1.
    """
    # Traverse from rightmost index to left
    for i in range(len(nums) - 1, -1, -1):
        if nums[i] == target:
            return i
    return -1

# Example usage:
nums = [1, 3, 5, 3, 2]
target = 3
result = find_last_occurrence(nums, target)
print("Last occurrence of", target, "is at index:", result)
```

- **Right-to-Left Traversal:**

The function iterates over the array starting from the last index down to 0. This guarantees that the first time the target is found (when traversing from right to left) is its last occurrence in the array.

- **Time Complexity:**

The traversal takes $O(n)$ time in the worst-case scenario (when the target is at the beginning of the array or not present), which meets the requirements.

Sorting The Array

When you receive an unsorted array and decide to sort it before applying the two-pointer technique, the overall time complexity is dominated by the sorting step.

- **Sorting:** This typically takes $O(n \log n)$ time.
- **Two Pointer Traversal:** Once sorted, scanning the array with two pointers takes $O(n)$ time.

Thus, the overall time complexity becomes:

$$O(n \log n) + O(n) = O(n \log n)$$

If the array is already sorted, then you only pay the $O(n)$ cost for the two-pointer traversal, without the additional $O(n \log n)$ sorting cost.

Index as a Hash Key

String Cheatsheet

A string is a sequence of characters. Many tips that apply to Arrays also apply to Strings.

Time complexity

A Strings is an array of characters, so the time complexities of basic string operations will closely resemble that of array operations.

| Operation | Big-O |
|-----------|--------|
| Access | $O(1)$ |
| Search | $O(n)$ |
| Insert | $O(n)$ |
| Remove | $O(n)$ |

Operations involving another String

Here we assume the other string is of length m .

| Operation | Big-O |
|---|------------|
| Find substring | $O(n.m)$ |
| Concatenating strings | $O(n + m)$ |
| Slice | $O(m)$ |
| Split (by token) | $O(n + m)$ |
| Strip (remove leading and trailing whitespaces) | $O(n)$ |

Things to look out for during interviews

- Ask about input character set and case sensitivity.

Corner cases

- Empty string
- String with 1 or 2 characters

- String with repeated characters
- Strings with only distinct character

02_Strings

Counting Characters in a String

String of unique characters

A neat trick to count the characters in a string of unique characters is to use a 26-bit bitmask to indicate which lower case latin characters are inside the string.

```
mask = 0
for c in word:
    mask |= (1 << (ord(c) - ord('a')))
```

■

To determine if two strings have common characters, perform `&` on the two bitmasks. If the result is non-zero, ie. `mask_a & mask_b > 0`, then the two strings have common characters.

String of Unique Characters

Counting characters

Often you will need to count the frequency of characters in a string. The most common way of doing that is by using a hash table/map in your language of choice.

If you need to keep a counter of characters, a common mistake is to say that the space complexity required for the counter is $O(n)$. The space required for a counter of a string of latin characters is $O(1)$ not $O(n)$. This is because the upper bound is the range of characters, which is usually a fixed constant of 26. The input set is just lowercase Latin characters.

Anagram

An anagram is word switch or word play. It is the result of rearranging the letters of a word or phrase to produce a new word or phrase, while using all the original letters only once.

In interviews, usually we are only bothered with words without spaces in them.

To determine if two strings are anagrams, there are a few approaches:

- Sorting both strings should produce the same resulting string.
 - This takes $O(n \cdot \log(n))$ time and $O(\log(n))$ space.
- If we map each character to a prime number and we multiply each mapped number together, anagrams should have the same multiple (prime factor decomposition).
 - This takes $O(n)$ time and $O(1)$ space. Examples: [Group Anagram](#)
- Frequency counting of characters will help to determine if two strings are anagrams.
 - This also takes $O(n)$ time and $O(1)$ space.

Palindrome

Palindrome

A palindrome is a word, phrase, number, or other sequence of characters which reads the same backward as forward, such as `madam` or `racecar`.

Here are ways to determine if a string is a palindrome:

- Reverse the string and it should be equal to itself.
- Have two pointers at the start and end of the string. Move the pointers inward till they meet. At every point in time, the characters at both pointers should match.

The order of characters within the string matters, so hash tables are usually not helpful.

When a question is about counting the number of palindromes, a common trick is to have two pointers that move outward, away from the middle.

- Note that palindromes can be even or odd length. For each middle pivot position, you need to check it twice - once that includes the character and once without the character.
- For substrings, you can terminate early once there is no match