Министерство науки и высшего образования Российской Федерации

Федеральное государственное автономное образовательное учреждение высшего образования

НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИТМО

(Университет ИТМО)

ОТЧЕТ ПО ЛАБОРАТОРНОЙ РАБОТЕ №3

по курсу «Алгоритмы и структуры данных»

Тема: БЫСТРАЯ СОРТИРОВКА, СОРТИРОВКИ ЗА ЛИНЕЙНОЕ ВРЕМЯ

Вариант 21

СОДЕРЖАНИЕ

[ЗАДАЧИ ПО ВАРИАНТУ 4](#_Toc181018147)

[1 Сортировка слиянием 4](#_Toc181018148)

[Задание 4](#_Toc181018149)

[Код 4](#_Toc181018150)

[Анализ кода 5](#_Toc181018151)

[Вывод 5](#_Toc181018152)

[2 Нахождение количества инверсий 6](#_Toc181018153)

[Задание 6](#_Toc181018154)

[Код 6](#_Toc181018155)

[Анализ кода 7](#_Toc181018156)

[Вывод 7](#_Toc181018157)

[3 Бинарный поиск 8](#_Toc181018158)

[Задание 8](#_Toc181018159)

[Код 8](#_Toc181018160)

[Анализ кода 8](#_Toc181018161)

[Вывод 9](#_Toc181018162)

[4 Поиск мажорирующего элемента 10](#_Toc181018163)

[Задание 10](#_Toc181018164)

[Код 1 10](#_Toc181018165)

[Анализ кода 1 10](#_Toc181018166)

[Код 2 11](#_Toc181018167)

[Анализ кода 2 11](#_Toc181018168)

[Вывод 11](#_Toc181018169)

[5 Поиск подмассива с максимальной суммой за линию 12](#_Toc181018170)

[Задание 12](#_Toc181018171)

[Код 12](#_Toc181018172)

[Анализ кода 12](#_Toc181018173)

[Вывод 13](#_Toc181018174)

[6 Умножение полиномов 14](#_Toc181018175)

[Задание 14](#_Toc181018176)

[Код 14](#_Toc181018177)

[Анализ кода 15](#_Toc181018178)

[Вывод 16](#_Toc181018179)

[ВЫВОД ПО ЛАБОРАТОРНОЙ РАБОТЕ 17](#_Toc181018180)

# ЗАДАЧИ ПО ВАРИАНТУ

## Быстрая сортировка

### Задание

Реализовать алгоритм быстрой сортировки

### Код

from typing import List, Tuple, TypeVar  
from random import randint  
  
  
T = TypeVar("T")  
  
  
def partition(lst: List[T], start: int, end: int, pivot: T) -> Tuple[int, int]:  
 end\_left = end\_mid = start  
 for i in range(start, end):  
 if lst[i] < pivot:  
 lst[i], lst[end\_mid], lst[end\_left] = lst[end\_mid], lst[end\_left], lst[i]  
 end\_left += 1  
 end\_mid += 1  
 elif lst[i] == pivot:  
 lst[end\_mid], lst[i] = lst[i], lst[end\_mid]  
 end\_mid += 1  
  
 return end\_left, end\_mid  
  
  
def quick\_sort(lst: List[T], start: int = 0, end: int = -1) -> None:  
 if end == -1:  
 end = len(lst)  
  
 if end - start < 2:  
 return  
  
 end\_left, end\_mid = partition(lst, start, end, lst[randint(start, end - 1)])  
 quick\_sort(lst, start, end\_left)  
 quick\_sort(lst, end\_mid, end)

### Анализ кода

1. **Функция** partition

Принимает подмассив от start до end, переменную pivot и разделяет массив на три части:

* Элементы, меньшие pivot (слева).
* Элементы, равные pivot (посередине).
* Элементы, большие pivot (в оставшейся части).

Проходит по массиву один раз, используя указатели end\_left и end\_mid для отслеживания текущих позиций, чтобы обменивать значения и сохранять относительный порядок элементов.

1. **Функция** quick\_sort**:**

Рекурсивно сортирует подмассивы. Вызов с диапазоном [start, end\_left) сортирует элементы, меньшие pivot. Вызов с [end\_mid, end) сортирует элементы, большие pivot.

Завершает работу, если в подмассиве менее двух элементов.

Время работы в среднем случае: O(n\*log ₂n), где n - длина списка. Время работы в худшем случае: O(n^2)

### Вывод

## Нахождение количества инверсий

### Задание

Найти количество инверсий в перестановке за O(n\*log ₂n)

### Код

import typing as tp  
  
  
def merge\_count\_inversions(list1: tp.List[int], list2: tp.List[int], target: tp.List[int], start\_index: int = 0) -> int:  
 *"""  
 Слияние list1 и list2, подсчет суммы по i от 0 до длины list1: количество элементов из list2, меньших list1[i]  
 :param list1: отсортированный список 1  
 :param list2: отсортированный список 2  
 :param target: список, в который будет помещён результат слияния списков 1 и 2  
 :param start\_index: индекс в списке target, откуда начнётся запись элементов  
 :return: сумма по i от 0 до длины list1: количество элементов из list2, меньших list1[i]  
 """* cnt = 0  
 cur1 = cur2 = 0 # текущие индексы в list1 и list2 соотв.  
 cur\_target = start\_index # текущий индекс в target  
 while cur1 < len(list1) and cur2 < len(list2):  
 if list1[cur1] <= list2[cur2]:  
 target[cur\_target] = list1[cur1]  
 cur1 += 1  
 cnt += cur2  
 else:  
 target[cur\_target] = list2[cur2]  
 cur2 += 1  
 cur\_target += 1  
  
  
 if cur1 == len(list1):  
 for i in range(cur2, len(list2)):  
 target[cur\_target] = list2[i]  
 cur\_target += 1  
 else:  
 for i in range(cur1, len(list1)):  
 target[cur\_target] = list1[i]  
 cur\_target += 1  
 cnt += (len(list1) - cur1) \* len(list2)  
  
 return cnt  
  
  
def merge\_sort\_count\_inversions(lst: list) -> int:  
 *"""  
 Нахождение количества инверсий  
 :param lst: список, в котором нужно найти количество инверсий (отсортируется в процессе)  
 :return: количество инверсий в lst  
 """* count\_inv = 0 # счетчик инверсий  
 len\_merging\_lists = 1 # длина сливающихся подсписков  
 while len\_merging\_lists < len(lst):  
 start\_index = 0 # начальный индекс пары сливающихся подсписков  
 while start\_index + len\_merging\_lists < len(lst):  
 if lst[start\_index + len\_merging\_lists - 1] <= lst[start\_index + len\_merging\_lists]:  
 start\_index += 2 \* len\_merging\_lists  
 continue  
  
 count\_inv += merge\_count\_inversions(lst[start\_index : start\_index + len\_merging\_lists],  
 lst[start\_index + len\_merging\_lists : min(len(lst), start\_index + 2 \* len\_merging\_lists)],  
 lst, start\_index)  
  
 start\_index += 2 \* len\_merging\_lists  
 len\_merging\_lists \*= 2  
  
 return count\_inv

### Анализ кода

Инверсии в массиве — это такие пары элементов, что i < j, а A[i] > A[j]. Подсчёт количества инверсий можно провести при помощи модифицированного алгоритма сортировки слиянием, используя тот факт, что количество инверсий списка равно сумме количества инверсий двух его частей и суммы, где каждое i-е слагаемое – количество элементов из правой части, таких, что они больше i-ого элемента из левой части. Первое и второе слагаемые находятся рекурсивно, а третье можно посчитать при слиянии. Причем, чтобы сделать эти вычисления быстрее, все-таки необходимо отсортировать части списка. Тогда, если при сравнении текущих элементов частей списков получается, что текущий элемент левой части меньше текущего из правой, можно утверждать, что все элементы правой, меньше текущего элемента из правой части (а таких элементов = индекс текущего правого), меньше текущего из левой.

Алгоритм состоит из:

1. Разделения списка на две части.
2. Рекурсивного подсчёта инверсий в каждой из частей.
3. Подсчёта инверсий, возникающих при слиянии двух частей.

**Оценка времени работы**: аналогично обычной сортировке слиянием

Время работы: O(n\*log ₂n), аналогично сортировке слиянием.

Затраты памяти: O(n) на временные массивы для слияния.

### Вывод

Встроенная модификация алгоритма сортировки слиянием позволяет одновременно с сортировкой подсчитывать инверсии.

## Бинарный поиск

### Задание

Реализовать алгоритм бинарного поиска элемента в списке.

### Код

import typing as tp  
  
  
def bin\_pow(lst: tp.List[int], value: int) -> int:  
 *"""  
 Поиск value в списке lst  
 :param lst: список, в котором будет производиться поиск  
 :param value: искомый элемент  
 :return: индекс искомого элемента или -1, если его нет  
 """* lst\_indexes = list(range(len(lst)))  
 lst\_indexes.sort(key=lambda index: lst[index])

left = 0  
 right = len(lst) - 1  
 while left <= right:  
 mid = left + (right - left) // 2  
 if lst[lst\_indexes[mid]] == value:  
 return lst\_indexes[mid]  
 if lst[lst\_indexes[mid]] < value:  
 left = mid + 1  
 else:  
 right = mid - 1  
  
 return -1

### Анализ кода

Бинарный поиск — это алгоритм поиска элемента в отсортированном списке. Он работает по принципу "разделяй и властвуй", на каждом шаге деля список пополам и проверяя, находится ли искомый элемент в левой или правой части.

Чтобы быстро выводить индекс элемента в неотсортированном списке, мы сортируем не сами элементы, а их индексы: индекс i является больше индекса j, если lst[i] больше lst[j]. Т.е. на позиции i хранится значение позиции в неотсортированном массиве, где стоял элемент, который в отсортированном массиве имеет индекс i.

Вместо рекурсии опять же использовался итеративный алгоритм, где роль разделения массива на две части играют два указателя, показывающие, с какой частью списка мы работаем. Алгоритм итеративного бинарного поиска начинается с установки начального значения границ поиска (left = 0, right = n-1). Затем определяется позиция mid как среднее между left и right. Элемент на позиции mid сравнивается с целевым значением. Если они равны, поиск завершен успешно. В противном случае, если элемент меньше целевого, left увеличивается, а если больше, right уменьшается. Процесс продолжается до тех пор, пока left не станет больше right, указывая на отсутствие элемента.

**Оценка времени работы**: аналогично обычной сортировке слиянием

Время работы: O(log ₂n), где n - это количество элементов в массиве. Это происходит потому, что каждый шаг поиска делит массив на две половины, сокращая область поиска примерно вдвое.

Затраты памяти: O(1) на переменные.

### Вывод

Бинпоиск – понятный (даже без использования рекурсий) и эффективный алгоритм.

## Поиск мажорирующего элемента

### Задание

Найти элемент, который встречается в списке более половины раз, если такой есть

### Код 1

import typing as tp  
  
  
def majority\_element\_recursion(lst: tp.List[int], start: int = 0, end: int = -1) -> tp.Tuple[tp.Optional[int], int]:  
 *"""  
 Поиск мажорирующего элемента на части lst от start до end (не включительно)  
 :param lst: список  
 :param start: начальный индекс списка  
 :param end: конечный индекс (end = len(lst) <-> end == -1)  
 :return: значение мажорирующего элемента (None, если его нет) и  
 количество раз, которое он встречается в списке (-1, если его нет)  
 """* if end == -1:  
 end = len(lst)  
  
 num\_elements = end - start  
 if num\_elements == 0:  
 return None, -1  
 if num\_elements == 1:  
 return lst[start], 1  
  
 mid = start + num\_elements // 2  
 majority\_left, num\_left = majority\_element\_recursion(lst, start, mid)  
 majority\_right, num\_right = majority\_element\_recursion(lst, mid, end)  
  
 if num\_left != -1:  
 for i in range(mid, end):  
 if lst[i] == majority\_left:  
 num\_left += 1  
 if num\_left > num\_elements // 2:  
 return majority\_left, num\_left  
  
 if num\_right != -1:  
 for i in range(start, mid):  
 if lst[i] == majority\_right:  
 num\_right += 1  
 if num\_right > num\_elements // 2:  
 return majority\_right, num\_right  
  
 return None, -1

### Анализ кода 1

Алгоритм реализован рекурсивно. Каждый раз он делит исходный список на две части, находит мажорирующий элемент для левой и правой частей, после чего проверяет, являются ли мажорирующие элементы частей (если они найдены) мажорирующими элементами на всем списке.

**Оценка времени работы**:

Время работы: O(n\*log ₂n), где n - это количество элементов в массиве.

Затраты памяти: O(log ₂n).

### Код 2

def majority\_element\_line(lst: tp.List[int]) -> tp.Tuple[tp.Optional[int], int]:  
 *"""  
 Поиск мажорирующего элемента в lst за линию  
 :param lst: список  
 :return: значение мажорирующего элемента (None, если его нет) и  
 количество раз, которое он встречается в списке (-1, если его нет)  
 """* num\_without\_pair = 0 # количество элементов которые пока без пары  
 candidate = None # Значение элементов без пары  
   
 for elem in lst:  
 if num\_without\_pair == 0:  
 candidate = elem  
 num\_without\_pair += 1  
 elif elem == candidate:  
 num\_without\_pair += 1  
 else:  
 num\_without\_pair -= 1  
  
 cnt = 0  
 for elem in lst:  
 if elem == candidate:  
 cnt += 1  
  
 return (candidate, cnt) if cnt > len(lst) // 2 else (None, -1)

### Анализ кода 2

Алгоритм основан на том факте, что если в списке есть мажорирующий элемент, то если поделить элементы на пары, такие что ни в какой паре не встречается 2 одинаковых элемента, без пары останутся элементы, значение которых и будет мажорирующим.

Таким образом, алгоритм предполагает, что в списке есть мажорирующий элемент, находит кандидата и проверяет, действительно ли он мажорирующий.

**Оценка времени работы**:

Время работы: O(n), где n - это количество элементов в массиве. Т.к. алгоритм просто проходится по списку два раза.

Затраты памяти: O(1).

### Вывод

Возможно, первая реализация интуитивно больше понятна, тем не менее она сильно уступает по времени и затратам памяти второму алгоритму.

## Поиск подмассива с максимальной суммой за линию

### Задание

Найти подмассив с максимальной суммой и вывести его сумму, начальный и конечный индексы.

### Код

import typing as tp  
  
  
def find\_max\_subarray(lst: tp.List[int]) -> tp.Tuple[int, int, int]:  
 *"""  
 Поиск максимального подмассива  
 :param lst: список  
 :return: сумма элементов максимального подмассива, стартовый индекс, конечный индекс  
 """* max\_subarray = lst[0] # сумма элементов максимального подмассива  
 left, right = 0, 1 # границы максимального подмассива  
  
  
 sm = 0 # сумма подмассива от 0 до текущего индекса  
 min\_sm = min\_sm\_len = 0 # подмассив от 0 до min\_sm\_len с минимальной суммой  
 for ind, elem in enumerate(lst):  
 sm += elem  
 if max\_subarray < sm - min\_sm:  
 max\_subarray = sm - min\_sm  
 left = min\_sm\_len  
 right = ind + 1  
  
 if min\_sm > sm:  
 min\_sm = sm  
 min\_sm\_len = ind + 1  
  
 return max\_subarray, left, right

### Анализ кода

Алгоритм основан на следующих фактах:

1. Зная сумму всех подмассивов, начинающихся с первого элемента, можно легко вычислить сумму любого подмассива. Т.е., чтобы узнать, какая сумма у подмассива lst[left : right] нужно из суммы подмассива lst[: right] вычесть сумму подмассива lst[: left].
2. Чтобы узнать сумму максимального подмассива, заканчивающегося элементом cur – 1, нужно знать сумму подмассива lst[: cur] и сумму минимального подмассива, начинающегося с первого элемента (заканчиваться он может максимум элементом cur - 2). Отнимая от первого значения второе, получаем исковый резуьтат.

Используя эти факты, алгоритм совершает следующие действия:

Считает текущую сумму подмассива lst[: cur] = lst[: cur - 1] + cur.

От этого значения отнимает сумму минимального подмассива, начинающегося с первого элемента. Таким образом получается значение максимального подмассива, заканчивающегося элементом cur-1.

Если полученная сумма больше всех подобных, ранее подсчитанных сумм, т.е. больше максимума из сумм подмассивов массива lst[:cur - 1], то максимальная сумма подмассивов массива [:cur] будет равна значению максимального подмассива, заканчивающегося элементом cur - 1. Иначе она останется равной максимальной сумме подмассивов массива lst[:cur - 1].

В конце обновляется минимум из сумм подмассивов, начинающихся первым элементом.

**Оценка времени работы**:

Время работы: O(n), где n - это количество элементов в массиве. Т.к. все происходит за один проход по списку.

Затраты памяти: O(1) на переменные.

### Вывод

Алгоритм Кадане, реализованный выше, - эффективный алгоритм нахождения максимального подмассива, который можно оптимизировать так, чтобы память практически не использовалась.

## Умножение полиномов

### Задание

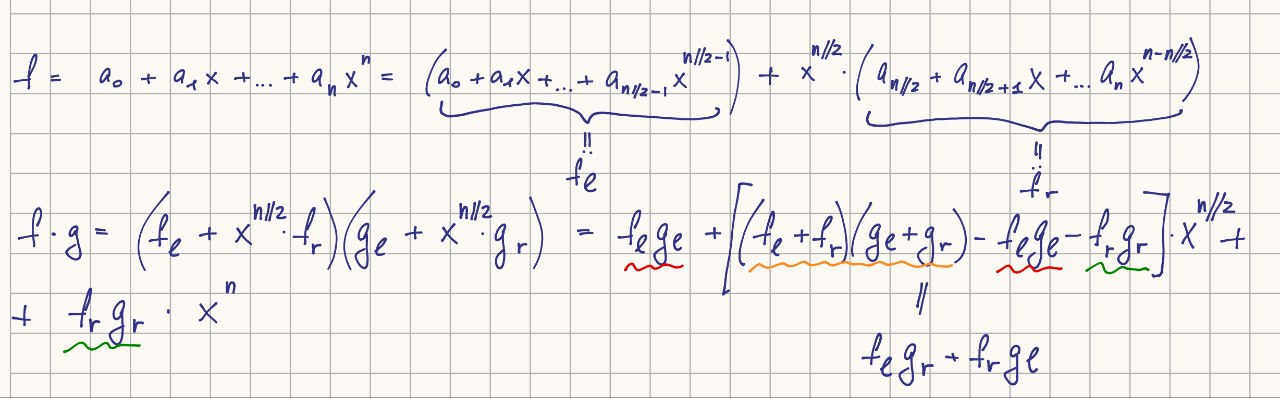
Реализовать алгоритм умножения полиномов менее чем за O(n^2)

### Код

def mult\_polynomials(f: tp.List[int], g: tp.List[int]) -> tp.List[int]:  
 *"""  
 Умножение полинома f на полином g (если степени полиномов разные, они дополняются нулями)  
 Полиномы представляются в виде:  
 f = [a0 + a1 \* x + a2 \* x^2 + ...], g = [b0 + b1 \* x + b2 \* x^2 + ...]  
 :param f: [a0, a1, a2, ...]  
 :param g: [b0, b1, b2, ...]  
 :return: полином степени deg(f) + deg(g) - 1 в аналогичном виде  
 """* if not f or not g:  
 raise ValueError("Empty polynomials")  
  
 if len(f) < len(g):  
 f.extend([0] \* (len(g) - len(f)))  
 elif len(f) > len(g):  
 g.extend([0] \* (len(f) - len(g)))  
  
  
 def mult\_polynomials\_recursion(f: tp.List[int], g: tp.List[int],  
 f\_start: int, f\_end: int,  
 g\_start: int, g\_end: int) -> tp.List[int]:  
 len\_f = f\_end - f\_start  
 len\_g = g\_end - g\_start  
 assert len\_f == len\_g  
  
 if len\_f == 1:  
 return [f[f\_start] \* g[g\_start]]  
  
 mid = len\_f // 2  
 f\_mid = f\_start + mid  
 g\_mid = g\_start + mid  
  
 f\_left\_g\_left = mult\_polynomials\_recursion(f, g, f\_start, f\_mid, g\_start, g\_mid)  
 f\_right\_g\_right = mult\_polynomials\_recursion(f, g, f\_mid, f\_end, g\_mid, g\_end)  
  
 f\_left\_plus\_f\_right = f[f\_mid:f\_end]  
 g\_left\_plus\_g\_right = g[g\_mid:g\_end]  
 for i in range(f\_start, f\_mid):  
 f\_left\_plus\_f\_right[i - f\_start] += f[i]  
 for i in range(g\_start, g\_mid):  
 g\_left\_plus\_g\_right[i - g\_start] += g[i]  
  
 sum\_f\_sum\_g = mult\_polynomials\_recursion(f\_left\_plus\_f\_right, g\_left\_plus\_g\_right,  
 0, f\_end - f\_mid,  
 0, g\_end - g\_mid)  
  
 result = [0] \* (len\_f \* 2 - 1)  
  
 for i in range(len(f\_left\_g\_left)):  
 result[i] += f\_left\_g\_left[i]  
 result[mid + i] -= f\_left\_g\_left[i]  
  
 for i in range(len(f\_right\_g\_right)):  
 result[mid \* 2 + i] += f\_right\_g\_right[i]  
 result[mid + i] -= f\_right\_g\_right[i]  
  
 for i in range(len(sum\_f\_sum\_g)):  
 result[mid + i] += sum\_f\_sum\_g[i]  
  
 return result  
  
 return mult\_polynomials\_recursion(f, g, 0, len(f), 0, len(g))

### Анализ кода

Алгоритм умножения полиномов, представленный в коде, основан на методе разделяй и властвуй. Этот подход разбивает полиномы на части, умножает их рекурсивно и собирает результат, используя стратегию, похожую на умножение чисел с разбиением по разрядам.

****

**Таким образом, можно рекурсивно вызвать три умножения полиномов (частей исходного), а потом собрать результат с учетом сдвига (степени x, на который умножается произведение)**

1. **Разбиение полиномов**:

Алгоритм делит каждый полином на две части:

* + Левая часть — это полиномы с меньшими степенями.
  + Правая часть — полиномы с большими степенями.

Например, если полином f(x) = a0 + a1⋅x + a2⋅x^2 + a3⋅x^3, то он будет разделен на:

* + Левую часть: f\_left(x) = a0 + a1⋅x
  + Правую часть: f\_right(x) = a2 + a3⋅x

1. **Рекурсивное умножение**:

После разбиения двух полиномов f и g на левые и правые части, алгоритм рекурсивно умножает:

* + Левую часть полинома​ f на левую часть полинома g.
  + Правую часть полинома f ​ на правую часть полинома g.

1. **Вычисление промежуточных сумм**:

Чтобы уменьшить количество операций умножения, алгоритм также складывает левую и правую части полиномов f и g (результаты копируются в новые списки) и умножает суммы. Это позволяет исключить лишние вычисления.

1. **Сборка результата**:

После рекурсивного умножения левых и правых частей, а также суммы левых и правых частей, результат собирается в один полином с поправками на позиции коэффициентов. Это делает алгоритм более эффективным по сравнению с наивным умножением, которое требует умножения каждого коэффициента одного полинома на каждый коэффициент другого.

**Оценка времени работы**:

Время работы: O(n^log ₂3), где n - длина списка.

### Вывод

Ужасный алгоритм с точки зрения использования памяти. Была попытка исправить это, но в рекурсии в любом случае нужно копирование списков (для вычисления списка суммы левой и правой частей полинома) перед рекурсивным вызовом. Тем не менее, алгоритм работает быстрее наивного.

# ВЫВОД ПО ЛАБОРАТОРНОЙ РАБОТЕ

Метод "разделяй и властвуй" широко используется в алгоритмах и имеет как преимущества, так и недостатки, которые зависят от природы задачи и особенностей реализации. Вот ключевые плюсы и минусы этого подхода:

**Преимущества**

1. Эффективность на больших задачах:

Алгоритмы, использующие метод разделяй и властвуй, часто имеют логарифмическую или линейно-логарифмическую сложность. Например, быстрая сортировка и сортировка слиянием работают за O(log ₂n), что делает их более эффективными по сравнению с квадратичными алгоритмами на больших входных данных.

2. Простота проектирования для рекурсивных задач:

Разделяя задачу на более мелкие подзадачи, можно сократить сложность логики каждой из них. Это упрощает разработку рекурсивных алгоритмов, таких как нахождение максимального подмассива, умножение матриц и вычисление дискретного преобразования Фурье.

**Недостатки**

1. Затраты на рекурсивные вызовы:

Рекурсивные реализации требуют дополнительной памяти для стека вызовов. Если глубина рекурсии велика, например, при неблагоприятных входных данных в быстрой сортировке, это может привести к переполнению стека. Тем не менее многие алгоритмы, использующие рекурсию можно переписать в итеративные, просто делать это иногда трудно, читаемость кода и его понимание ухудшаются.

2. Избыточное копирование данных:

В некоторых алгоритмах приходится копировать подмассивы или подстроки для обработки отдельных частей, что приводит к увеличению затрат по памяти и времени. Тем не менее некоторые алгоритмы вполне реально оптимизировать в этом плане, но в таком случае, код становится менее читаемым, подключается больше параметров.

Метод "разделяй и властвуй" подходит для задач, где большие объемы данных могут быть разбиты на независимые и простые подзадачи. Несмотря на его преимущества для повышения производительности и возможности параллелизации, он требует тщательной реализации, чтобы избежать затрат на копирование данных и учитывать потенциальные проблемы с глубокой рекурсией.