**Введение**

Олимпиадное программирование состоит из двух частей – проектирования алгоритмов и реализации алгоритмов.

Проектирование алгоритмов. По сути своей, олимпиадное программирование – это придумывание эффективных алгоритмов решения корректно поставленных вычислительных задач. Для проектирования алгоритмов необходимы навыки в решении задач и знание математики. Зачастую решение появляется в результате сочетания хорошо известных методов и новых идей.

Важную роль в олимпиадном программировании играет математика. На самом деле четких границ между проектированием алгоритмов и математикой не существует.

Реализация алгоритмов. В олимпиадном программировании решение задачи оценивается путем проверки реализованного алгоритма на ряде тестов. Поэтому придумать алгоритм недостаточно, его еще нужно корректно реализовать, для чего требуется умение программировать. Олимпиадное программирование сильно отличается от традиционной программной инженерии: программы короткие (несколько сотен строк – уже редкость), писать их нужно быстро, а сопровождение после соревнования не требуется.

В настоящее время на соревнованиях по программированию популярнее всего языки C++, Python и Java. Например, среди 3000 лучших участников Google Code Jam 2017 79% писали на C++, 16% – на Python и 8% – на Java.

На данном сайте Вы сможете ознакомиться с необходимой теорией для языка Python для успешного написания олимпиад в области информатики, а также проверить полученные знания на практике, решением предоставленных задач.

**Тема 1. Эффективность**

Во время своей работы программы используют различные структуры данных и алгоритмы, в связи с чем обладают разной эффективностью и скоростью решения задачи. Дать оценку оптимальности решения, реализованного в программе, поможет понятие вычислительной сложности алгоритмов.

1.1. Основные понятия

Вычислительная сложность (алгоритмическая сложность) - понятие, обозначающее функцию зависимости объема работы алгоритма от размера обрабатываемых данных.

Вычислительная сложность пытается ответить на центральный вопрос разработки алгоритмов: как изменится время исполнения и объем занятой памяти в зависимости от размера входных данных?. С помощью вычислительной сложности также появляется возможность классификации алгоритмов согласно их производительности.

В качестве показателей вычислительной сложности алгоритма выступают:

1. Временная сложность (время выполнения).

Временная сложность алгоритма - это функция от размера входных данных, равная количеству элементарных операций, проделываемых алгоритмом для решения экземпляра задачи указанного размера.

Временная сложность алгоритма зачастую может быть определена точно, однако в большинстве случаев искать точное ее значение бессмысленно, т.к. работа алгоритма зависит от ряда факторов, например, скорости процессора, набора его инструкций и т.д.

1. Асимптотическая сложность.

Асимптотическая сложность оценивает сложность работы алгоритма с использованием асимптотического анализа.

Алгоритм с меньшей асимптотической сложностью является более эффективным для всех входных данных.

1.2. Асимптотические нотации

Асимптотическая сложность алгоритма описывается соответствующей нотацией:

1. О-нотация, O («О»-большое): описывает верхнюю границу времени (время выполнения «не более, чем…»);
2. Омега-нотация, Ω («Омега»-большое): описывает нижнюю границу времени (время выполнения «не менее, чем…»).

Например, T(n)=O(N2) говорит о том, что алгоритм имеет квадратичное время выполнения относительно размера входных данных в качестве верхней оценки («О большое от эн квадрат»).

Каждая оценка при этом может быть:

1. наилучшая: минимальная временная оценка;
2. наихудшая: максимальная временная оценка;
3. средняя: средняя временная оценка.

При оценке, как правило, указывается наихудшая оценка.

Допустим, имеется задача поиска элемента в массиве. При полном переборе слева направо:

1. наилучшая оценка: O(1), если искомый элемент окажется в начале списка;
2. наихудшая оценка: O(N), если искомый элемент окажется в конце списка;
3. средняя оценка: O(N/2)=O(N)

1.3. Верхняя оценка и O-нотация

Наиболее часто используемой оценкой сложности алгоритма является верхняя (наихудшая) оценка, которая обычно выражается с использованием нотации O-большое.

Выделяют следующие основные категории алгоритмической сложности в O-нотации:

1. Постоянное время: O(1)

* Время выполнения не зависит от количества элементов во входном наборе данных.
* Пример: операции присваивания, сложения, взятия элемента списка по индексу и др.

1. Линейное время: O(N)

* Время выполнения пропорционально количеству элементов в коллекции.
* Пример: найти имя в телефонной книге простым перелистыванием, почистить ковер пылесосом и т.д.

1. Логарифмическое время: O(logN)

* Время выполнения пропорционально логарифму от количества элементов в коллекции.
* Пример: найти имя в телефонной книге (используя двоичный поиск).

1. Линейно-логарифмическое время: O(NlogN)

* Время выполнения больше чем, линейное, но меньше квадратичного.
* Пример: обработка N телефонных справочников двоичным поиском.

1. Квадратичное время: O(N2)

* Время выполнения пропорционально квадрату количества элементов в коллекции.
* Пример: вложенные циклы (сортировка, перебор делителей и т.д.)

1.4. Оценка сложности алгоритмов

Для оценки вычислительной сложности алгоритмов необходимо знать и учитывать сложности:

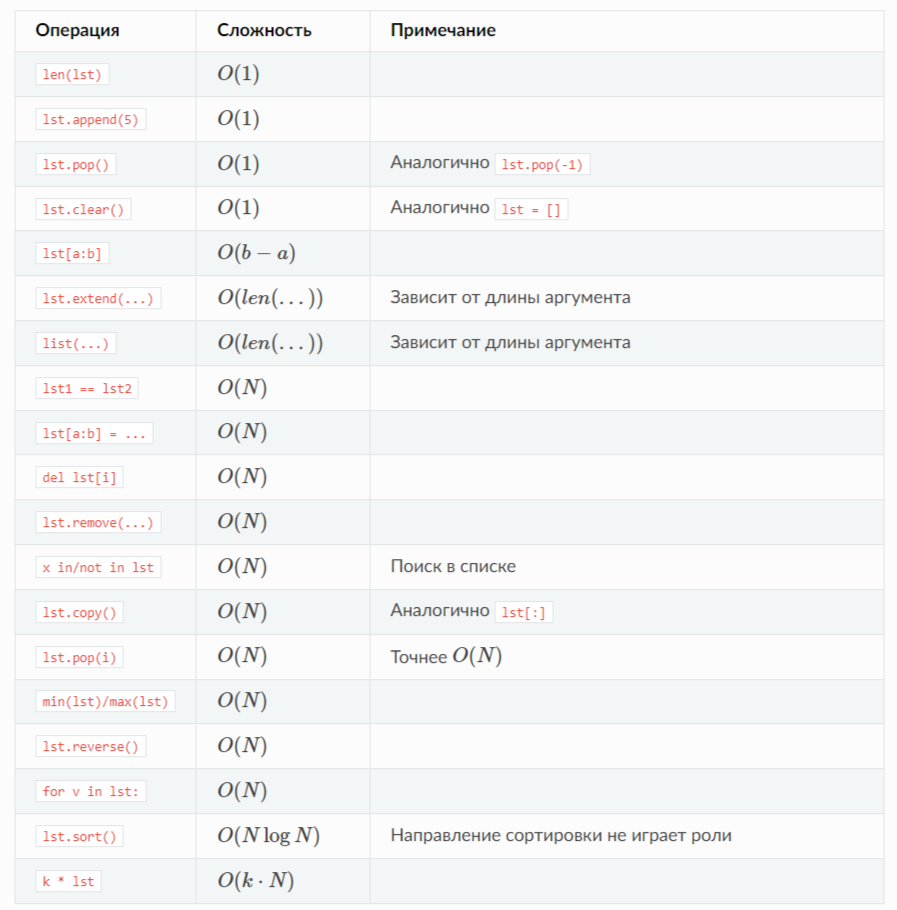
* используемых структур данных;
* совокупности различных операций.

1.5. Операции над структурами данных

В Python имеются коллекции (структуры данных), операции над которыми имеют определенную сложность.

1.5.1. Список и кортеж

Большинство операций со списком/кортежем имеют сложность O(N).



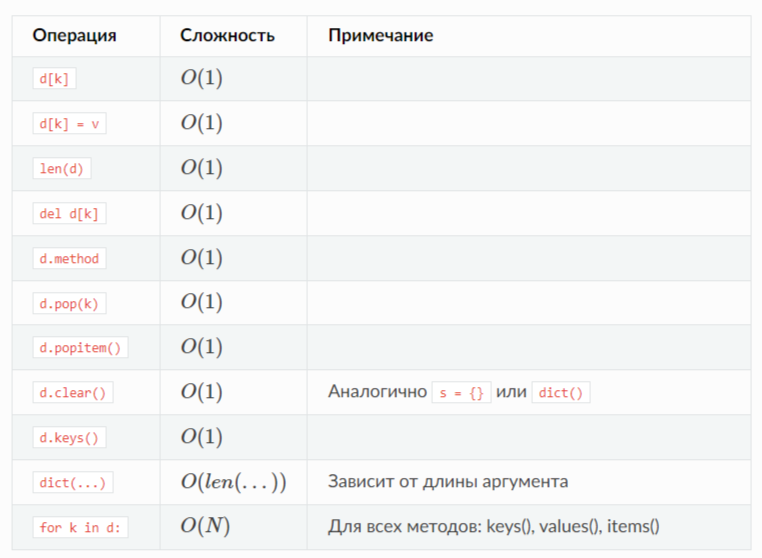
1.5.2. Множество

По сравнению со списком/кортежем множества большую часть операций выполняют со сложностью O(1).



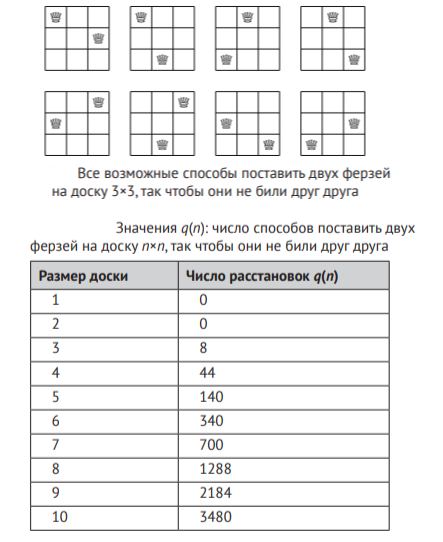
1.5.3. Словарь

Большинство операций словарей имеет сложность O(1).



1.6. Задача о двух ферзях

Разберём задачу: сколькими способами можно поставить на доску n×n двух ферзей, так чтобы они не били друг друга. На рис. показано, что на доске 3×3 это можно сделать 8 способами. Обозначим q(n) количество расстановок на доске n×n. Например, q(3) = 8, а в табл. приведены значения q(n) для 1 ≤ n ≤ 10.



Конечно, задачу можно решить по-простому: перебрать все возможные расстановки двух ферзей и подсчитать те, в которых ферзи не бьют друг друга. Временная сложность такого алгоритма O(n4 ), поскольку позицию первого ферзя можно выбрать n2 способами, а затем поставить второго ферзя n2 − 1 способами.

Поскольку число расстановок растет очень быстро, алгоритм их подсчета по одиночке будет работать слишком медленно для больших n. Следовательно, для создания эффективного алгоритма нужно придумать способ подсчета расстановок группами. Полезно отметить, что очень просто подсчитать, сколько клеток бьет один ферзь (видим это на следующем рисунке). Во-первых, он всегда бьет n − 1 клеток по горизонтали и n − 1 клеток по вертикали. Далее по обеим диагоналям ферзь бьет d − 1 клеток, где d – число клеток на диагонали. С учетом всего этого для вычисления числа клеток, в которые можно поставить второго ферзя, нужно время O(1), так что мы получаем алгоритм с временной сложностью O(n2).



К задаче можно подойти и по-другому, попытавшись найти рекуррентную формулу для подсчета расстановок, т. е. ответить на вопрос: как, зная q(n), вычислить q(n + 1)?

Чтобы получить рекурсивное решение, обратим внимание на последние горизонталь и вертикаль доски n×n (на представленном рисунке). Во-первых, число расстановок, в которых на последней горизонтали и на последней вертикали нет ферзей, равно q(n − 1). Во-вторых, общее число позиций ферзя на последней горизонтали и на последней вертикали равно 2n − 1. Находящийся в этих позициях ферзь бьет 3(n − 1) клеток, поэтому для второго ферзя остается n2 − 3(n − 1) – 1 позиций. Наконец, существует (n − 1)(n − 2) расстановок, в которых оба ферзя находятся на последней горизонтали или на последней вертикали. Поскольку эти расстановки посчитаны дважды, их количество надо вычесть. Объединив все вместе, получаем рекуррентную формулу:

q(n) = q(n − 1) + (2n − 1)(n2 − 3(n − 1) − 1) − (n − 1)(n − 2) = q(n − 1) + 2(n − 1)2 (n − 2)

и вместе с ней решение сложности O(n).



Однако же существует и замкнутая формула для этой функции:

q(n) = n4/2 - 5n3/3 + 3n2/2 - n/3,

которую можно доказать по индукции, пользуясь рекуррентной формулой. А раз так, то задачу можно решить за время O(1).

**Тема 2. Битовые операции: AND, OR, NOT, XOR - таблицы значений.**

Как известно, минимальной единицей измерения информации является **бит**, который хранит одно из 2-х значений: 0 (**False**, ложь) либо 1 (**True**, истина). Таким образом, битовая ячейка может одновременно находиться лишь в одном из двух возможных состояний.

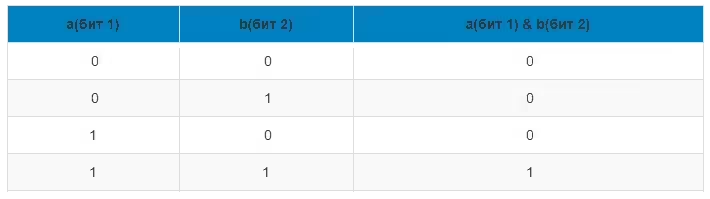
Для манипуляций с битами используют определённые операции — **логические или булевые**. Они могут применяться к любому биту, вне зависимости от того, какое у него значение — ноль или единица. Что же, давайте посмотрим на примеры использования трёх основных логических операций.

**Логическая операция AND (и)**

**AND** обозначается знаком &.

Оператор AND выполняется с 2-мя битами, возьмём, к примеру, a и b. Результат выполнения операции AND равен 1, если a и b равняются 1. В остальных случаях результат равен 0. Например, с помощью AND вы можете узнать, чётное число или нет.

Посмотрите на таблицу истинности операции AND:



**Логическая операция OR (ИЛИ)**

Обозначается знаком |.

Оператор **OR** также выполняется с 2-мя битами (a и b). Результат равен 0, если a и b равны 0, иначе он равен 1. Смотрим таблицу истинности.

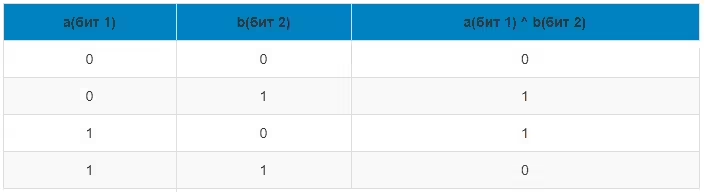


**Логическая операция XOR (исключающее ИЛИ)**

Оператор XOR обозначается ^.

**XOR** выполняется с 2-мя битами (a и b). Результат выполнения операции XOR (**исключающее ИЛИ**) равен 1, когда один из битов b или a равен 1. В остальных ситуациях результат применения оператора XOR равен 0.

Таблица истинности логической операции для XOR (исключающее ИЛИ) выглядит так:



Используя XOR (исключающее ИЛИ), вы можете поменять значения 2-х переменных одинакового типа данных, не используя временную переменную. А ещё, посредством XOR можно зашифровать текст, например:



Согласен, XOR — далеко не самый надёжный метод шифрования, но это не значит, что его нельзя сделать частью какого-либо шифровального алгоритма.

**Логическая операция NOT (НЕ)**

Это побитовое отрицание, поэтому выполняется с одним битом и обозначается ~.

Результат зависит от состояния бита. Если он в нулевом состоянии, то итог операции — единица и наоборот. Всё предельно просто.



Свойства операции XOR:

1. XOR и 0: x^0 = x
2. XOR с одинаковыми аргументами: x^x = 0
3. Коммутативность: x^y = y^x

Эти 4 логические операции следует запомнить в первую очередь, т. к. с их помощью можно получить практически любой возможный результат. Также существуют такие операции, как <<(побитовый сдвиг влево) и >>(побитовый сдвиг вправо).

**Тема 3. Проверка на простоту, разложение на множители. Наибольший общий делитель и наименьшее общее кратное: алгоритм Евклида. Признак Паскаля. Расширенный алгоритм Евклида.**

Простое число - это натуральное, целое положительное число n, которое делится только на единицу и на себя.

Пример рабочего алгоритма, определяющего число на простоту:

def IsPrime(a):

if a <= 1: return False

for d in range(2, a):

if a % d == 0:

return False

return True

Алгоритм разложения числа на простые множители:

def Prime\_fac(n):

i = 2

prime\_fac = []

while i \*\* 2 <= n:

while n % i == 0:

prime\_fac.append(i)

n = n / i

i = i + 1

if n > 1:

prime\_fac.append(n)

return prime\_fac

Наибольший общий делитель (НОД):

A, B **=** [int(x) **for** x **in** input().split()]

a, b **=** A, B

**while** b !**=** 0:

a, b **=** b, a**%**b

print (A**\***B**//**(A**\***B**//**a))

Наименьшее общее кратное (НОК):

A,B **=** [int(x) **for** x **in** input().split()]

a, b **=** A, B

**while** b!**=** 0:

a, b **=** b, a**%**b

print (A**\***B**//**a)

Алгоритм Евклида — это алгоритм, который используется для нахождения наибольшего делителя двух целых чисел.

def gcd\_rem\_division(num1, num2):

while num1 != 0 and num2 != 0:

if num1 >= num2:

num1 %= num2

else:

num2 %= num1

return num1 or num2

Треугольник Паскаля:

def PrintPasTriangle(rows):

row = [1]

for i in range(rows):

print(row)

row = [sum(x) for x in zip([0] + row, row + [0])]

Расширенный алгоритм Евклида:

def gcd\_extended(num1, num2):

if num1 == 0:

return (num2, 0, 1)

else:

div, x, y = gcd\_extended(num2 % num1, num1)

return (div, y - (num2 // num1) \* x, x)

Проверить данный алгоритм можно следующим образом:

a = gcd\_extended(426, 334)

print(f”Делитель равен {a[0]}, x = {a[1]}, y = {a[2]}”)

Делитель равен 2, x = 69, y = -88

Rак работает алгоритм?

Сначала проверяется, равно ли первое число нулю, если это так, то второе число является делителем, а коэффициенты равны 0 и 1, так как «num1 \* x + num2 \* y = y» в том случае, если y = 1, а левое произведение равно нулю.

Функция возвращает три числа: делитель, коэффициент x и коэффициент y. Для её реализации используется рекурсия, делитель получается тем же образом, что и в классическом рекурсивным алгоритме, а коэффициенты рекурсивно вычисляются по формулам:

x = y — (num2 // num1) \* x

y = x

**Тема 4. Линейный поиск**

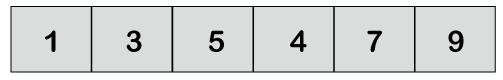
Линейный поиск – это метод поиска элементов в списке. Его еще называют последовательным поиском. Это простейший алгоритм поиска, поскольку он ищет желаемый элемент последовательно.

Он сравнивает каждый элемент со значением, которое мы ищем. Если оба совпадают, элемент найден, и алгоритм возвращает позицию индекса ключа.

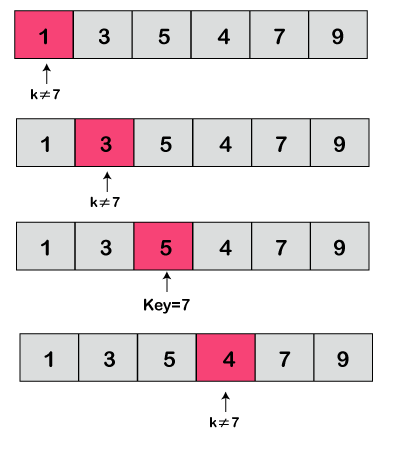
Концепция линейного поиска

Давайте разберёмся в следующих шагах, чтобы найти элемент key = 7 в данном списке.

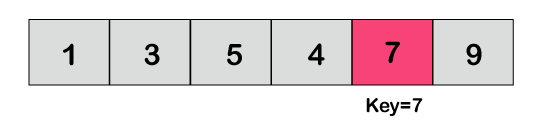
Шаг – 1: Начните поиск с первого элемента и проверьте ключ = 7 с каждым элементом списка x.



Шаг – 2: Если элемент найден, вернуть индексную позицию ключа.



Шаг – 3: Если элемент не найден, возвращаемого элемента нет.



Алгоритм решения данной задачи:

def linear\_Search(list1, n, key):

for i in range(0, n):

if (list1[i] == key):

return i

return -1

list1 = [1 ,3, 5, 4, 7, 9]

key = 7

n = len(list1)

res = linear\_Search(list1, n, key)

If res == -1:

print("Элемент не найден")

else:

print("Элемент найден по индексу ", res)

**Тема 5. Квадратичные сортировки (выбор, вставки, пузырек). Сортировка подсчетом. Быстрая сортировка Хоара.**

**Сортировка выбором**

Одна из наиболее часто возникающих в программировании задач — задача о сортировке элементов массива (списка). Постановка задачи — дан список элементов A, которые можно сравнивать (например, чисел, строк, кортежей и т. д.). Необходимо переставить элементы списка местами так, чтобы было выполнено условие A[i]<=A[i+1] для всех пар соседних элементов. Например, если был дан список [4, 1, 2, 4, 2, 3], то отсортированный список будет иметь вид [1, 2, 2, 3, 4, 4]. Такой порядок сортировки называется сортировкой по неубыванию элементов (но чаще используют не вполне точный термин «сортировка в порядке возрастания»). Если заменить условие на A[i] >= A[i+1], то получится сортировка в порядке невозрастания (убывания). Для сортировки списков придумано много различных алгоритмов.

Один из наиболее простых алгоритмов — сортировка выбором. Идея алгоритма следующая. Сначала выберем в списке наименьший элемент и поставим его на место с индексом 0 в списке (в начало списка). Потом среди всех оставшихся элементов выберем наименьший и поставим его на место с индексом 1. Затем выберем наименьший среди элементов элементов, начиная с третьего, и поставим его на место c индексом 2 и т. д.

Таким образом, в этой сортировке два вложенных цикла. Внешний цикл осуществляется по переменной i начиная с 0. При этом все элементы списка до элемента с индексом i (то есть A[:i]) есть наименьшие элементы списка, упорядоченные по неубыванию.

Теперь выберем среди элементов списка A[i:] элемент с наименьшим значением и поменяем его местами с элементом с индексом i.

def SelectionSort(A):

for i in range(0, len(A) - 1):

# Среди элементов A[i:] выбираем наименьший

# Сохраняем его индекс в переменной min\_idx

min\_idx = i

for j in range(i + 1, len(A)):

if A[j] < A[min\_idx]:

min\_idx = j

# Теперь поставим A[min\_idx] на место A[i]

A[i], A[min\_idx] = A[min\_idx], A[i]

Можно модифицировать алгоритм — не сохранять индекс наименьшего из просмотренных элементов, а при просмотре элементов в срезе A[i:] обменивать очередной элемент A[j] местами с A[i], если A[j]<A[i]:

def SelectionSort(A):

for i in range(0, len(A) - 1):

for j in range(i + 1, len(A)):

if A[j] < A[i]:

A[i], A[j] = A[j], A[i]

Посчитаем сложность этого алгоритма. Пусть список содержит n элементов. Сначала нужно найти минимум среди n элементов списка, что потребует n операций. Потом нужно найти наименьший из n-1 элемента, на это нужно n-1 операция. Потом нужно n-2 операции и т. д. Общее число операций равно

n + (n - 1) + (n - 2) + ... + 1 = n(n + 1)/2 = O(n2)

Таким образом, сортировка выбором — квадратичный алгоритм, время его работы пропорционально квадрату от размера списка.

**Сортировка вставками**

Сортировка вставками использует похожий инвариант: первый элементы списка, то есть срез A[:i] уже отсортирован. По-иному устроен алгоритм добавления i-го элемента к уже отсортированной части. Здесь берется элемент A[i] и добавляется к уже отсортированной части списка. Например, пусть i = 5 и срез A[:i] = [1, 4, 6, 8, 8], а значение A[i] == 5. Тогда элемент A[i] == 5 нужно поставить после элемента A[ 1] == 4, а все элементы, которые больше 5 сдвинуть вправо на 1. Получится cрез A[:i + 1] = [1, 4, 5, 6, 8, 8]. Таким образом, при вставке элемента A[i] в срез A[:i] так, чтобы в результате получился упорядоченный срез, все элементы, которые больше A[i] будут двигаться вправо на одну позицию. А в освободившуюся позицию и будет вставлен элемент A[i].

При этом значение A[i] нужно сохранить в переменной, т. к. на место элемента A[i], возможно, будет записан элемент A[i – 1].

Получаем следующий алгоритм:

def InsertionSort(A):

for i in range(1, len(A)):

# В new\_elem сохранили значение A[i]

new\_elem = A[i]

# Начиная с элемента A[i - 1]

j = i - 1

# все элементы, которые больше new\_elem

while j >= 0 and A[j] > new\_elem:

# сдвигаем вправо на 1

A[j + 1] = A[j]

j -= 1

# На свободное место записываем new\_elem

A[j + 1] = new\_elem

Посчитаем сложность алгоритма сортировки вставками. Следует отметить, что если массив уже упорядочен, то все элементы останутся на своем месте и вложенный цикл не будет выполнен ни разу. В этом случае сложность алгоритма сортировки вставками — линейная, т. е. . Аналогично, если массив «почти упорядочен», то есть для превращения его в упорядоченный нужно поменять местами несколько соседних или близких элементов, то сложность также будет линейной.

Но если массив упорядочен в обратном порядке, например, каждый элемент больше следующего, а необходимо добиться обратного порядка, то каждый элемент будет перемещаться максимально влево, т. е. до самой крайней позиции. В этом случае количество выполняемых перемещений будет равно

1 + 2 + ... + (n - 1) + n = n(n + 1)/2 = O(n2)

Итак, мы видим, что сложность алгоритма сортировки вставками сильно зависит от того, «насколько хорошо» отсортирован исходный список. В лучшем случае время работы — линейно, в худшем случае — квадратично. Что же происходит «в среднем», когда массив заполнен числами в случайном порядке?

В этом случае математическое ожидание количества перемещений элементов будет равно половине от числа перемещений в худшем случае (каждый элемент в среднем будет перемещаться не до самого начала списка, а только до середины этого пути), то есть математическое ожидание числа перемещений будет равно n(n + 1)/4 = O(n2). То есть в среднем этот алгоритм также имеет квадратичную сложность.

**Сортировка пузырьком**

Алгоритм сортировки пузырьком построен на простой идее. Пусть есть два соседних элемента, которые стоят в неправильном порядке, то есть A[i] > A[i + 1]. Поменяем их местами. Оказывается, такой операции уже достаточно, чтобы отсортировать массив — достаточно повторять такую операцию до тех пор, пока есть соседние неправильно упорядоченные элементы. Но необходимо еще организовать процесс так, чтобы он завершился.

Пройдем по всему списку слева направо. Если есть два неправильно упорядоченных элемента — переставим их. В результате самый большой элемент списка «всплывет» в его конец — станет последним элементом. Повторим этот проход еще раз — второй по величине элемент списка «всплывет» в конец, остановившись перед наибольшим элементов. За следующий проход мы можем установить на место третий элемент и т. д. При этом последние, уже «всплывшие» наибольшие элементы не нужно затрагивать алгоритмом сортировки. Получим следующий алгоритм:

def BubbleSort(A):

for j in range(len(A) - 1, 0, -1):

for i in range(0, j):

if A[i] > A[i + 1]:

A[i], A[i + 1] = A[i + 1], A[i]

Легко видеть, что сложность этого алгоритма также будет O(n2).

Что произойдет, если запустить сортировки пузырьком на уже остортированном списке? Ни одной перестановки не будет произведено, но алгоритм все равно выполнит O(n2) операций. Хотя уже после первого прохода вложенного цикла можно понять, что список уже упорядочен, если ни одной перестановки не было сделано. Это позволяет соптимизировать алгоритм сортировки — закончим его, если во внутреннем цикле не было выполнено ни одной перестановки. Для этого заведем переменную IsNotOrdered, которая будет равна True, если список не упорядочен. Перед проходом внутреннего цикла мы будем устанавливать IsNotOrdered=False (априори считаем, что список уже упорядочен), но если обнаруживается пара неупорядоченных элементов, то выполняется присваивание IsNotOrdered=True. Внешний цикл выполняется, пока переменная IsNotOrdered принимает значение True.

def BubbleSort(A):

j = len(A) - 1

IsNotOrdered = True

while IsNotOrdered:

IsNotOrdered = False

for i in range(0, j):

if A[i] > A[i + 1]:

A[i], A[i + 1] = A[i + 1], A[i]

IsNotOrdered = True

j -= 1

Такой алгоритм также будет работать за линейное время на почти упорядоченных массивах (довольно скоро массив упорядочится и внешний цикл закончится).

Есть различные алгоритмы, построенные на основе пузырьковой сортировке. Например, в шейкерной сортировке внутренний цикл проходится поочередно слева направо, затем справа налево. А в сортировке Шелла элементы переставляются не соседние элементы, а элементы, отстоящие на большее расстояние, что позволяет быстрее перемещать элементы по списку, выполняя сразу несколько шагов.

Быстрая сортировка Хоара.

Этот алгоритм, чаще называемый просто «быстрая сортировка» (англ. Quicksort) придуман английским ученым Чарльзом Хоаром в 1960 году.

Во многом идея быстрой сортировки такая же, как у алгоритма сортировки слиянием. Выберем некоторый элемент q, называемый барьерным элементом. Разобьем массив на две части, переупорядочив его элементы. В первой части соберем элементы, меньшие или равные q, а во второй части — большие или равные q. Теперь достаточно отсортировать обе части, после чего выполнить их конкатенацию безо всякого дополнительного слияния.

Простая реализация быстрой сортировки Хоара выглядит так:

def QuickSort(A):

if len(A) <= 1:

return A

else:

q = random.choice(A)

L = []

M = []

R = []

for elem in A:

if elem < q:

L.append(elem)

elif elem > q:

R.append(elem)

else:

M.append(elem)

return QuickSort(L) + M + QuickSort(R)

В данном примере в списке L собираются элементы, меньшие q, в списке R — большие q, а в списке M — равные q. Разделение на три списка, а не на два используется для того, чтобы алгоритм не зацикливался, например, в случае, когда в списке остались только равные элементы. Барьерный элемент q выбирается случайным образом из списка при помощи функции choice из модуля random.

Тот же алгоритм можно записать еще проще в «функциональном» стиле:

def QuickSort(A):

if len(A) <= 1:

return A

else:

q = random.choice(A)

L = [elem for elem in A if elem < q]

M = [q] \* A.count(q)

R = [elem for elem in A if elem > q]

return QuickSort(L) + M + QuickSort(R)

Однако, такая реализация алгоритма, как и сортировка слиянием, требует O(n) дополнительной памяти. Возможна реализация алгоритма сортировки слиянием без использования дополнительной памяти:

def QuickSort(A, l, r):

if l >= r:

return

else:

q = random.choice(A[l:r + 1])

i = l

j = r

while i <= j:

while A[i] < q:

i += 1

while A[j] > q:

j -= 1

if i <= j:

A[i], A[j] = A[j], A[i]

i += 1

j -= 1

QuickSort(A, l, j)

QuickSort(A, i, r)

Эта реализация не возвращает никакого значения, а сортирует элементы списка A «на месте», то есть модифицирует переданный список A. Два дополнительный параметра l и r указывают на номер первого и последнего элемента того фрагмента списка, который нужно отсортировать (включая эти элементы), то есть элемент сортирует срез A[l:r+1]. Для сортировки всего списка A необходимо вызвать QuickSort(A, 0, len(A) – 1).

Если l >= r, то ничего сортировать не нужно, срез пустой или содержит один элемент. Иначе случайным образом выбирается барьерный элемент. Далее заводятся два указателя i = l и j = r. Затем элементы списка переставляются так, чтобы элементы, которые меньше или равны q оказывались слева от указателя i, а те, которые больше или равны q оказывались справа от j. После окончания распределения рекурсивно сортируются две получившиеся части списка.

Данная реализация не требует дополнительной памяти (за исключением памяти, необходимой для организации рекурсии).

Асимптотика алгоритма

Сложность алгоритма быстрой сортировки Хоара зависит от метода выбора барьерного элемента. В лучшем случае при каждом выборе барьерного элемента должен выбираться медианный элемент массива. Но поиск медианного элемента — сложная задача, её нельзя решить быстро. Если выбрать первый элемент фрагмента списка A[l] или последний A[r], то если список A уже упорядочен, сложность алгоритма будет O(n2), так как на каждом рекурсивном вызове от большей части списка будет отделяться всего один элемент.

Поэтому в алгоритме быстрой сортировки Хоара, как правило, в качестве барьерного элемента выбирается случайный элемент списка. Тогда алгоритм становится вероятностным — время его работы зависит от того, каким будет случайно выбранный элемент. Возможна (но крайне маловероятна) ситуация, когда всегда будет выбираться наименьший элемент, и в этом случае алгоритм будет работать за O(n2).

В теории вероятностей доказывается, чти при случайном выборе элемента списка и разбиении его на две части, размер большей из двух получившихся частей будет в среднем равен 3**π** / 4. В этом случае глубина рекурсии в среднем будет составлять порядка log n, а средняя сложность алгоритма быстрой сортировки Хоара — O(n log n).

Тема 6.