

Лабораторная работа I.

Асимптотическая сложность

Зелинский Виктор

8 марта 2024

Цель работы: прямыми измерениями удостоверится, что асимптотическая сложность алгоритма совпадает с теоритической

Поиск

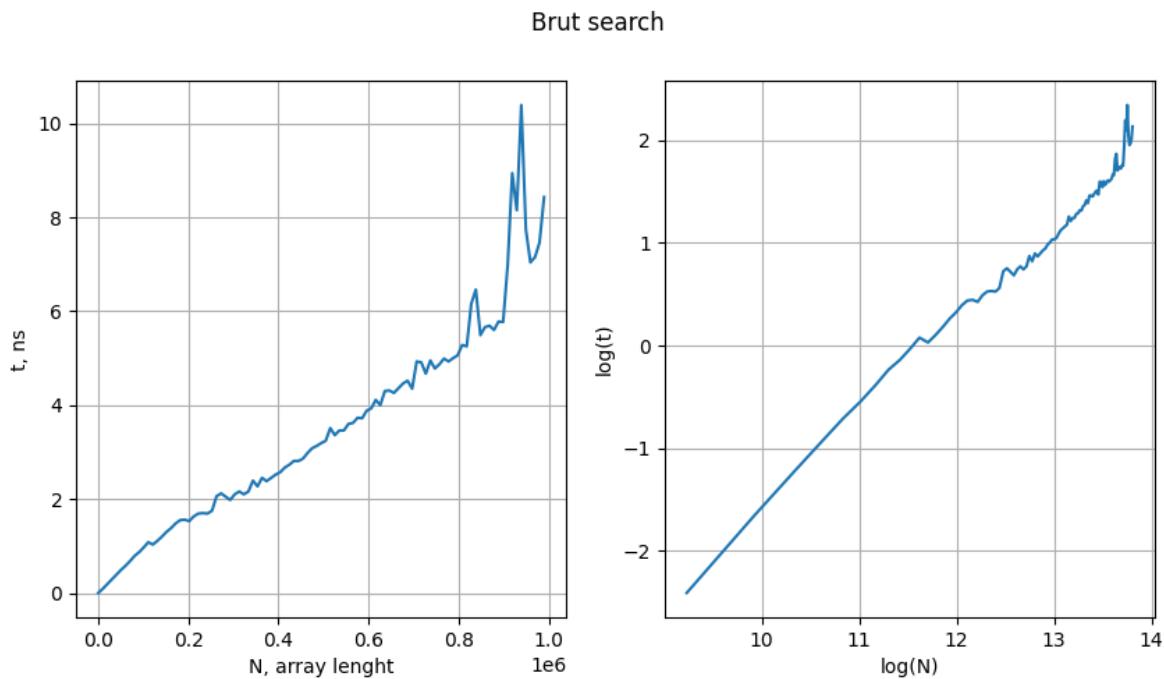


Рис. 1: График зависимости времени от длины массива (справа в логарифмическом масштабе)

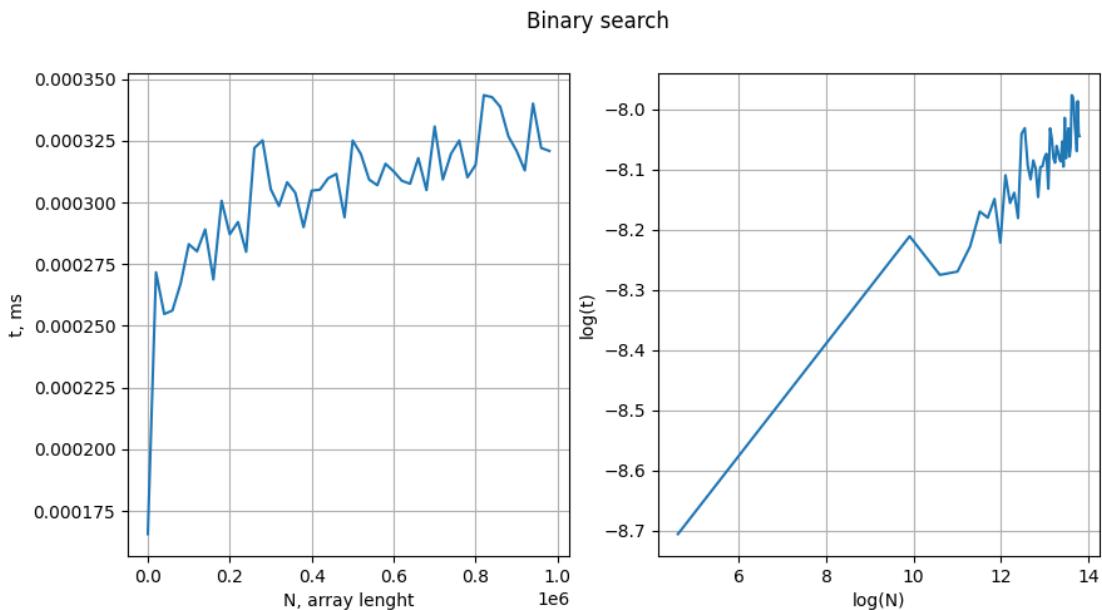


Рис. 2: График зависимости времени от длины массива (справа в логарифмическом масштабе)

Сумма двух

По графику (Рис. 3) можно определить, что сложность поиска в отсортированном массиве $O(N)$, а в неотсортированном $O(N^2)$, что совпадает с теоритическими расчетами.

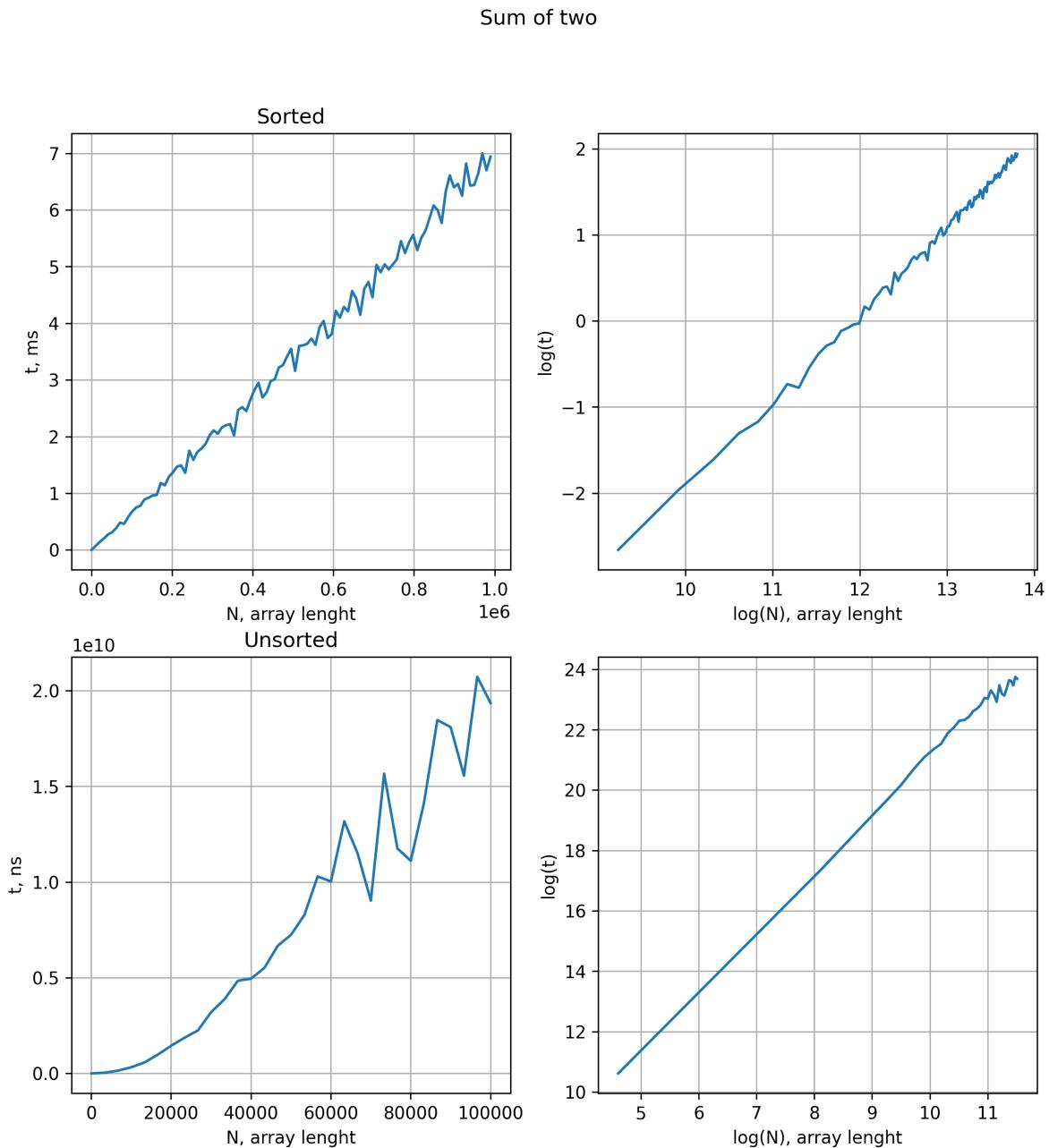


Рис. 3: График зависимости времени от длины массива (справа в логарифмическом масштабе)

Часто используемый элемент

Как бы грустно это не было, оптимизированные стратегии не дают прироста в скорости ни на равномерном распределении запросов, ни на биномиальном.

Smart search

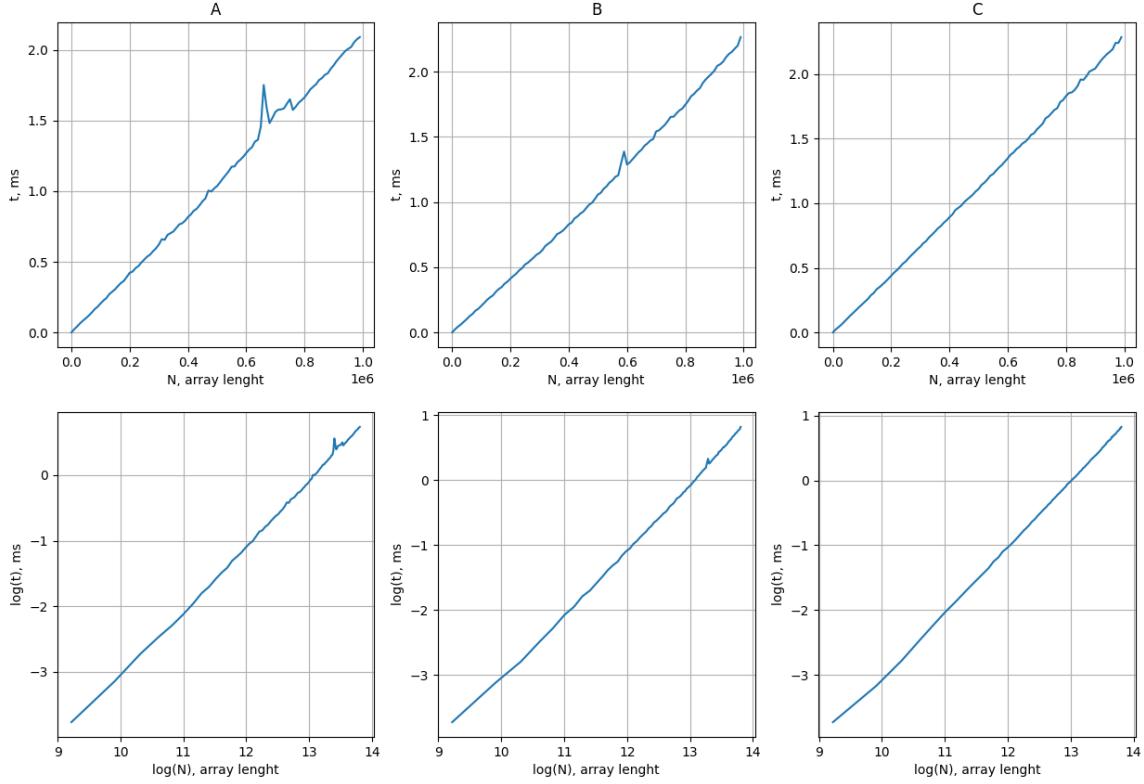


Рис. 4: График зависимости времени от длины массива (снизу в логарифмическом масштабе)

Действительно, теория говорит тоже самое. Оценим сложность при равномерном распределении запросов:

$$\frac{1 + 2 + \dots + N}{N} = \frac{N(N+1)}{2N} = \frac{N+1}{2} = O(N)$$

А теперь оценим при биномиальном распределении запросов

$$\frac{1 \cdot C_{2N}^N + 2C_{2N}^{N-1} + 3C_{2N}^{N+1} + \dots + NC_{2N}^0}{2N} < \frac{2N \cdot C_{2N}^N + 2NC_{2N}^{N-1} + 2NC_{2N}^{N+1} + \dots + 2NC_{2N}^0}{2N} = 2N$$

С другой стороны по транснеравенству

$$\frac{1 \cdot C_{2N}^N + 2C_{2N}^{N-1} + 3C_{2N}^{N+1} + \dots + NC_{2N}^0}{2N} > \frac{1}{2^N} \sum_{n=0}^{2N} nC_{2N}^n = N \frac{2^{N-1} - 1}{2^N}$$

таким образом, это $O(N)$

по собственной инициативе

Я написал проги, которые делает замер времени для одного размера массива, и запускал для разных в многопоточке через питон. Получил вот такие красивые графики.(Рис. 6)

Выводы

1. асимптотическая сложность может зависеть от входных данных, но не обязательно.

Smart search with binomial distribution

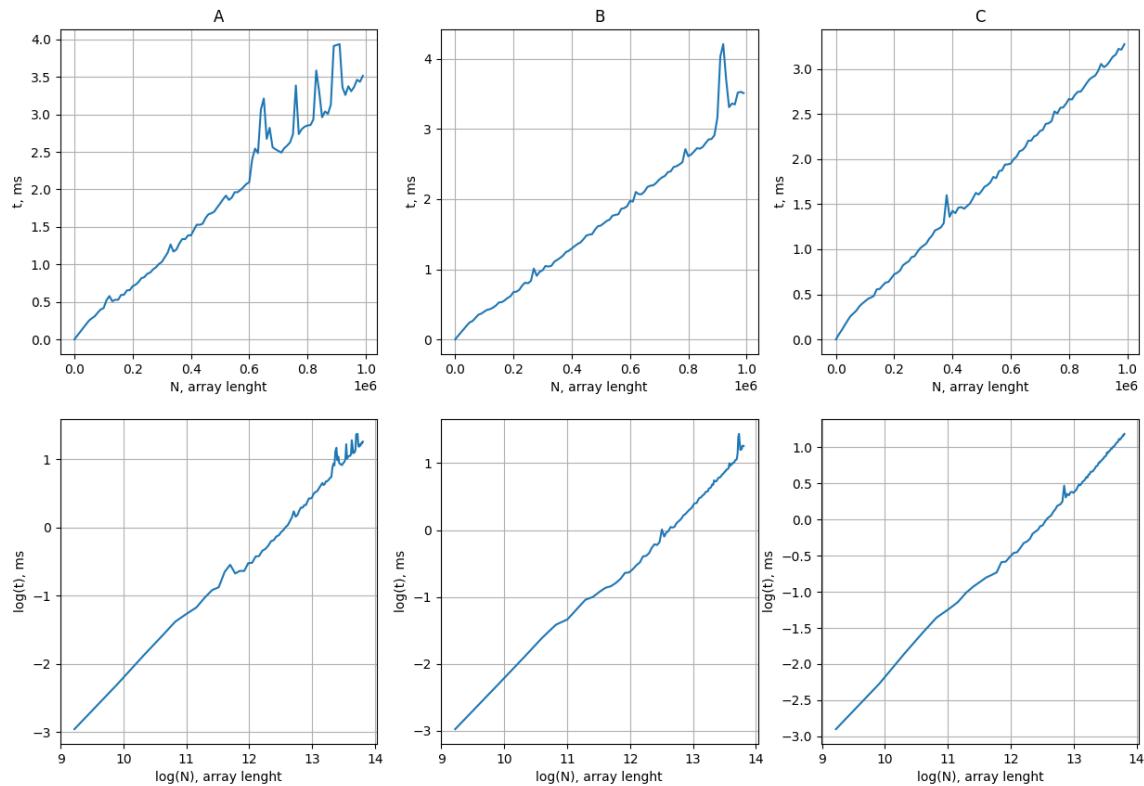


Рис. 5: График зависимости времени от длины массива (снизу в логарифмическом масштабе)

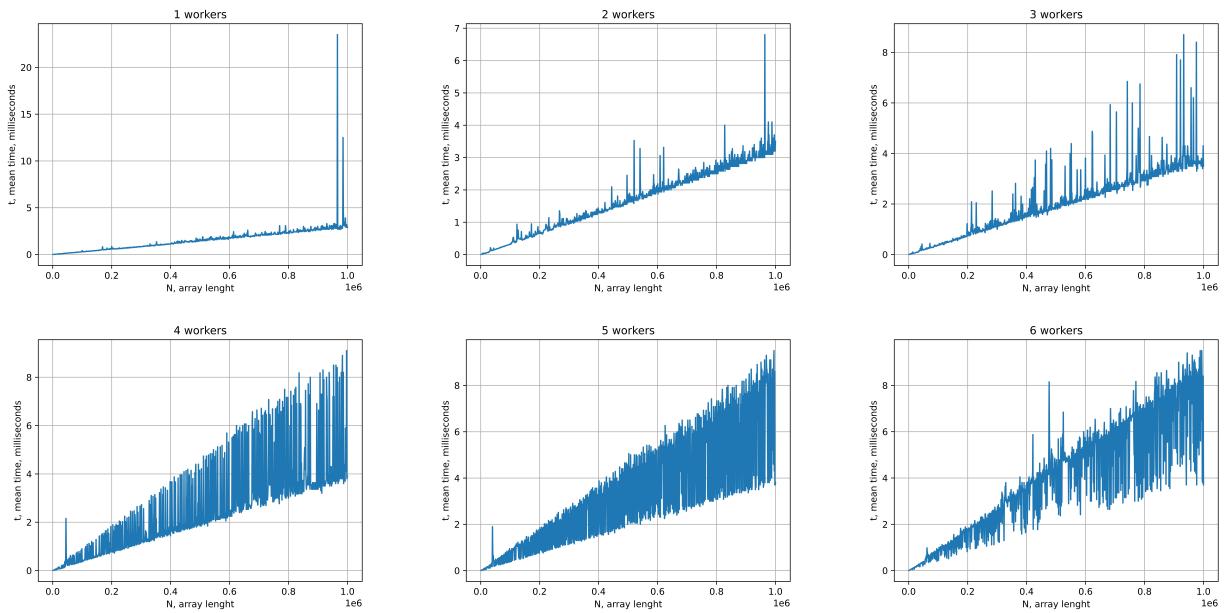


Рис. 6: График зависимости времени от длины массива (снизу в логарифмическом масштабе)

2. Весь этот бред про "для достаточно больших N " на самом деле работает и для вообразимых величин, но иногда если алгоритм не работает с большими входными данными и в оптимизированном виде есть много предварительных вычислений, лучше оставить старый.