# Міністерство освіти і науки України Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського" Факультет інформатики та обчислювальної техніки

# Кафедра ІІІІ

## Звіт

з лабораторної роботи № 2 з дисципліни «Алгоритми та структури даних 2. Структури даних»

"Метод декомпозиції. Пошук інверсій"

Виконав(ла) <u>ІП-15 Плугатирьов Дмитро Валерійович</u> (шифр, прізвище, ім'я, по батькові)

**Перевірив**— Соколовський Владислав Володимирович

— (прізвище, ім'я, по батькові)

### Завдання

Існує веб сервіс, який надає своїм користувачам можливість перегляду фільмів онлайн. Періодично система надає нові рекомендації користувачам — які фільми, що їх користувач ще не дивився, можливо будуть йому або їй цікаві.

В основі рекомендаційного алгоритму лежить ідея, що користувачі, які подивились однакові фільми та також оцінили їх схожим чином, мають схожі смаки. Наприклад, нехай є два користувача: Аліса та Богдан. Обидва вони переглянули наступні фільми: "Зоряні війни", "Гравітація", "Пірати карибського моря", "Володар перснів", "Матриця".

Спочатку система просить користувачів оцінити ці фільми і розташувати їх у порядку вподобання, іншими словами — створити власний хіт-парад. Так Аліса розташувала вказані фільми у порядку від найбільш до найменш вподобаного: "Пірати карибського моря", "Володар перснів", "Матриця", "Гравітація", "Зоряні війни". Хіт-парад Богдана: "Зоряні війни", "Володар перснів", "Гравітація", "Матриця", "Пірати карибського моря".

Після цього система може надати кількісну оцінку наскільки схожими  $\epsilon$  смаки двох користувачів. Для цього використовується алгоритм підрахунку інверсій поміж двома масивами.

Нехай A[1..n] — масив з n чисел. Якщо i < j та A[i] > A[j], то пара (i, j) — інверсія в A.

Щоб звести задачу порівняння двох хіт-парадів до задачі підрахунку інверсій у нашому прикладі, побудуємо два масиви A та B. Масив A = [1, 2, 3, 4, 5]. Масив B будується наступним чином: елементом B[j]  $\epsilon$  число, яке відповідає позиції фільму в хіт-параді Богдана, який в хіт-параді Аліси посідав місце j. Наприклад, j = 1 у хіт-параді Аліси відповідає фільму "Пірати карибського моря". Цей фільм в списку Богдана стоїть на позиції 5, тому B[1] = 5. Загалом отримуємо масив B = [5, 2, 4, 3, 1].

Масив B = [5, 2, 4, 3, 1] має наступні інверсії (вказуються індекси елементів, а не їх значення): (1,2), (1,3), (1,4), (1,5), (2,5), (3,4), (3,5), (4,5).

Загалом 8 інверсій. І це число вказує наскільки сильно відрізняється список вподобань Аліси від списку вподобань Богдана. Ми порахували віддаленість списку Аліси від списку Богдана. Якщо порахувати цю відстань в іншому напрямку, то чи буде вона такою самою? Тобто визначити кількість інверсій в списку Аліси по відношенню до списку Богдана.

Сервіс перегляду фільмів онлайн має базу даних D вподобань користувачів. Ця база  $\epsilon$  матрицею.

Рядки цієї матриці відповідають користувачам, а стовпці — фільмам. Її розмірність u\*m, де u — це кількість користувачів, m — кількість фільмів. Кожний елемент матриці D[i, j] вказує на позицію фільму j в списку вподобань користувача i. Для спрощення припускаємо, що всі користувачі переглянули всі фільми.

Тепер щоб визначити наскільки подібні смаки деякого користувача х до смаків інших користувачів, система попарно порівнює списки вподобань х та всіх інших користувачів і не дорівнює х: за вказаним вище принципом підраховується кількість інверсій у масиві D[x] відносно масиву D[i].

Визначене число інверсій буде кількісною оцінкою наскільки смаки х  $\epsilon$  близькими до смаків кожного і — чим менше значення цього числа, тим більш подібними  $\epsilon$  смаки двох користувачів.

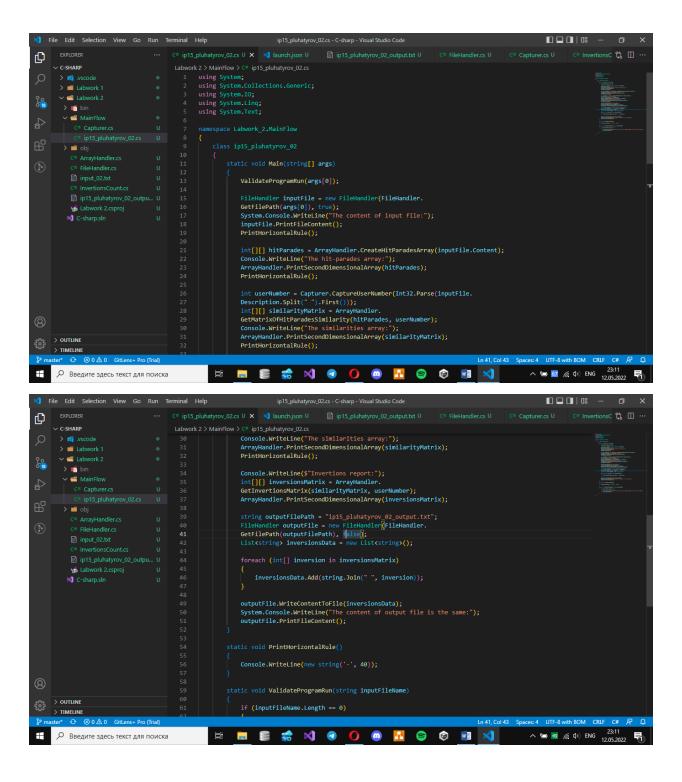
# Формальна постановка задачі

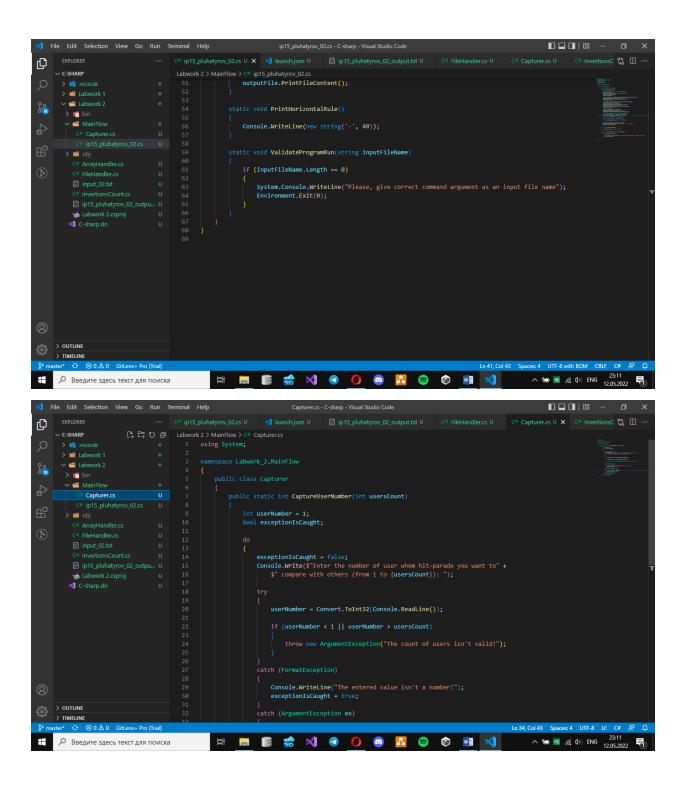
За допомогою методу декомпозиції розробити алгоритм, який буде розв'язувати наступну задачу.

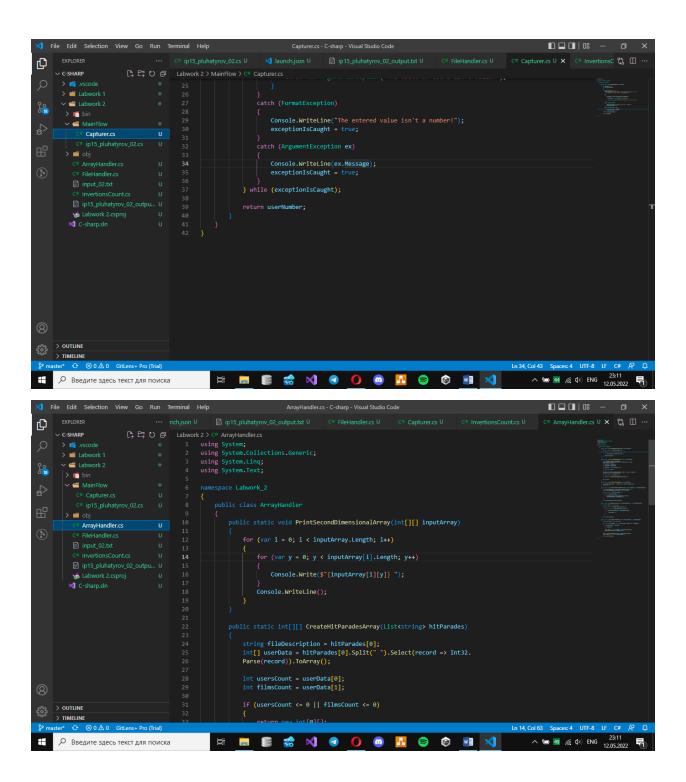
Вхідні дані. Матриця D натуральних чисел розмірності u\*m, де u — ці кількість користувачів, m — кількість фільмів. Кожний елемент матриці D[i, j] вказує на позицію фільму j в списку вподобань користувача i. Іншим вхідним елементом  $\epsilon$  х — номер користувача, з яким будуть порівнюватись всі інші користувачі.

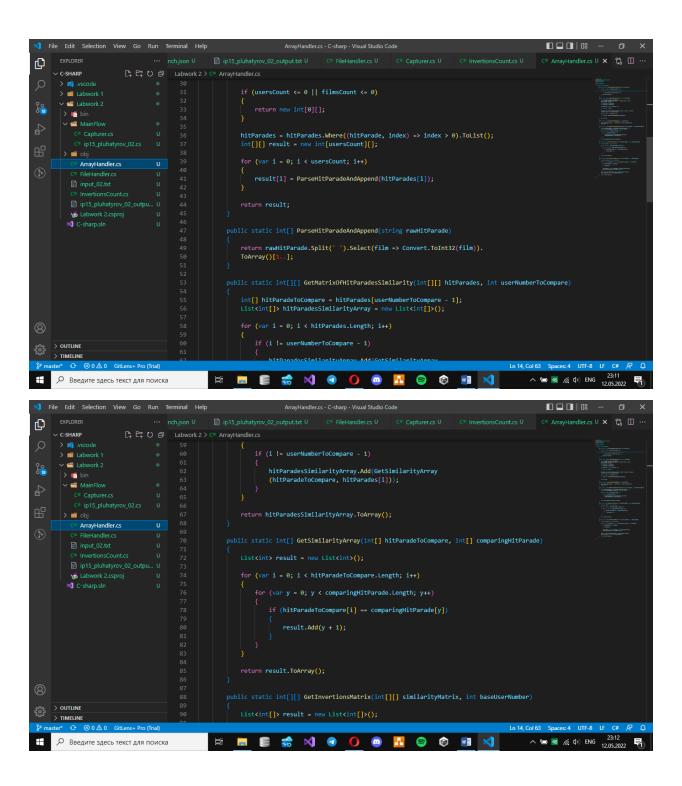
Вихідні дані. Список з впорядкованих за зростанням другого елементу пар (i, c), де і — номер користувача, с — число, яке вказує на степінь схожості вподобань користувачів х та с (кількість інверсій).

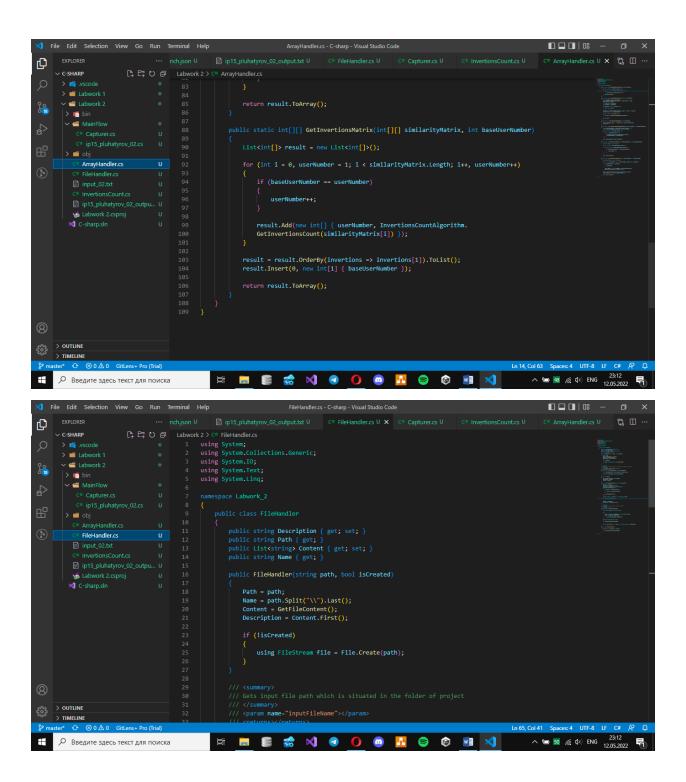
### Код програми

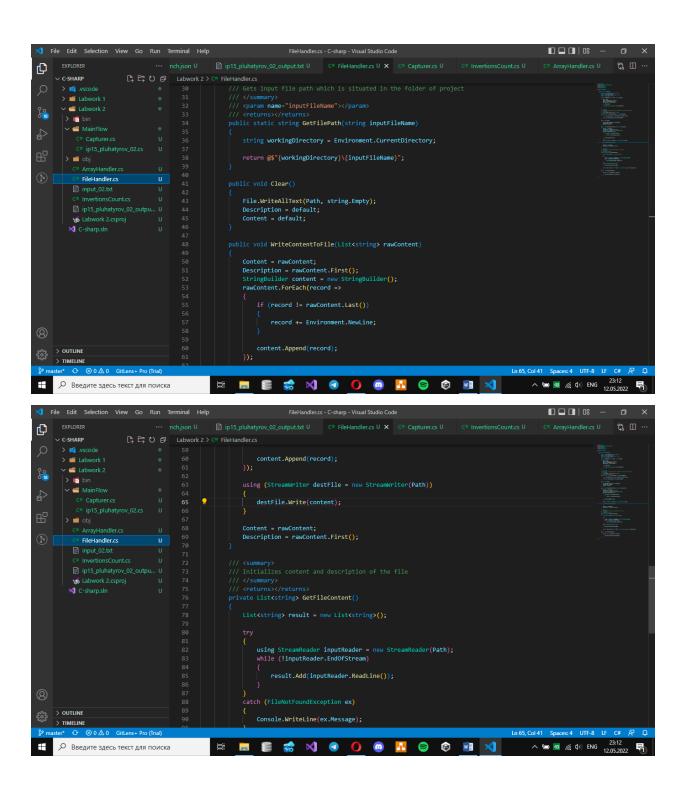


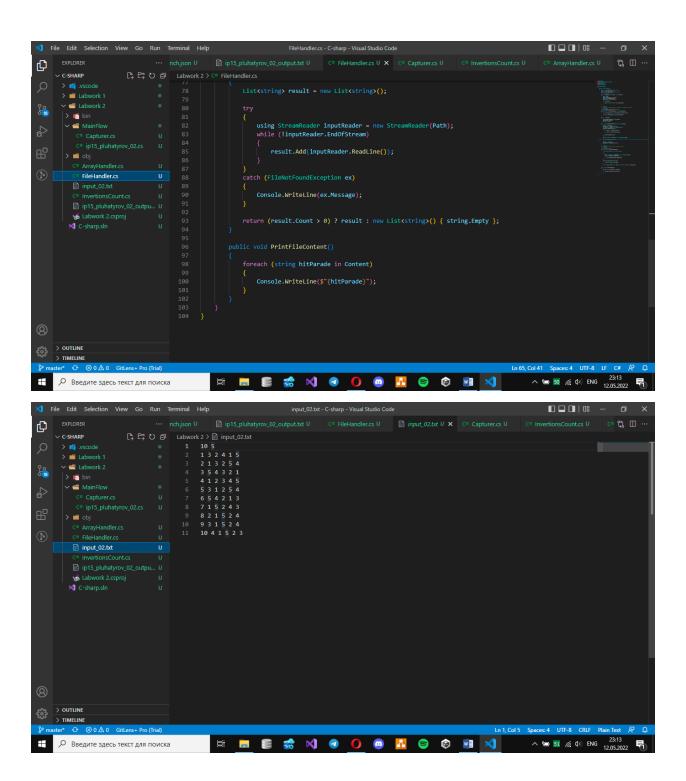


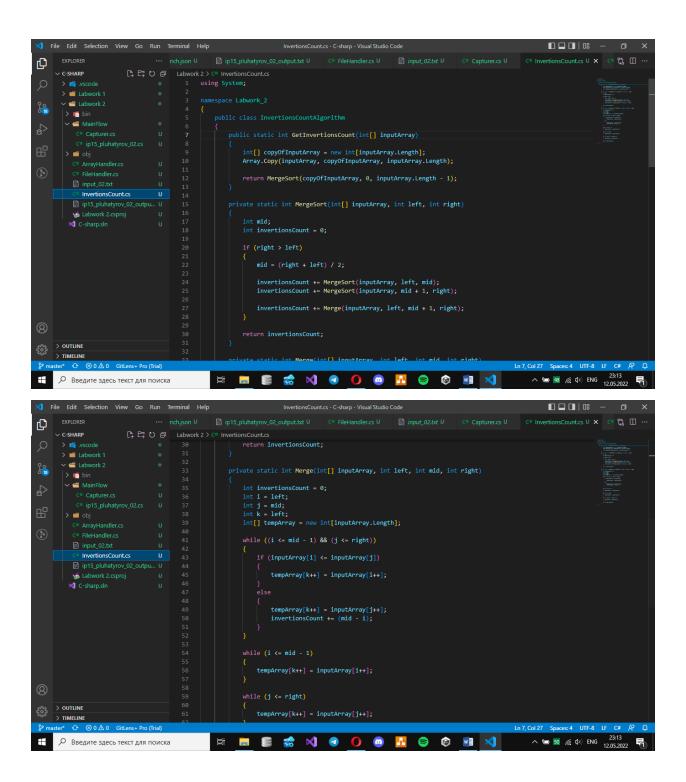


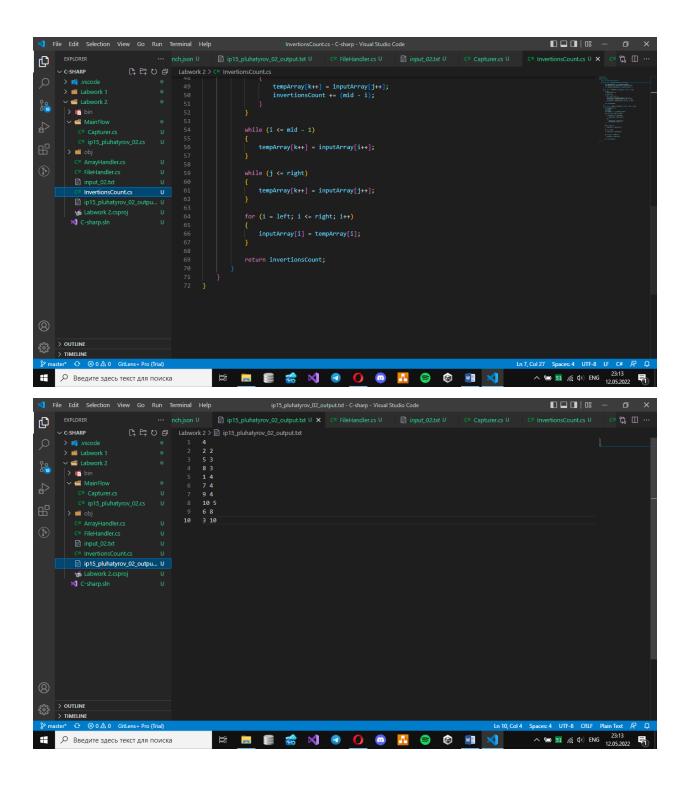












### Висновок

В цій лабораторній роботі я виконав порівняння хіт-парадів користувачів, підрахував кількість інверсій у кожній утвореній послідовності порівняння. Дані я брав із вхідного файлу, назву якого вводив в якості аргументу командного рядка перед запуском програми. Порівняння хіт-параду одного користувача з іншими, записується у вихідний файл. Для підрахунку кількості інверсій у послідовностях було використано алгоритм Merge sort.