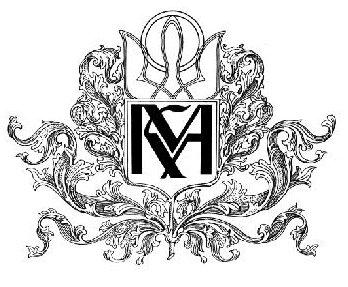
Національний університет “Києво-Могилянська академія”

Факультет інформатики

****

Звіт до лабораторної роботи

з курсу програмне забезпечення інтелектуальних систем

“Аналіз еволюції бінарних ланцюжків.”

Варіант 2

Виконали:

студенти МП СА-1

Матвеєва Марія та   
Степанюк Станіслав

Київ-2019

Зміст

1. Постановка задачі. 3
   1. Опис задачі 3

2. Завершення роботи 4

Висновок 5

# Постановка задачі.

## Опис задачі

**Мета**: Реалізувати генетичний алгоритм для випадку генераційного типу репродукції.

**Схема алгоритму**

0. Кодування розв’язків

1. Ініціалізація

2. Моделювання еволюційного процесу (репродукція)

2.1. Оцінювання. Якщо виконується умова зупинки, перехід на п.3

2.2. Відбір

2.3. Застосування оператора мутації

2.4. Перехід на п.2.1

3. Завершення роботи

***Розв’язки кодуються наступним чином:***

Бінарні ланцюжки довжини *l*, *l*  - параметр.

Розглянути випадки: *l*=10; *l*=20; *l*=80; *l*=100; *l*=200; *l*=800; *l*=1000; *l*=2000; *l*=8000; *l*=16000.

**Ініціалізація.** Всі особини початкової популяції - ланцюжки, що складаються з усіх 0: 00…0;

**Кількість особин в популяції**.

*N* – змінюється (зростає).

Розглянути варіанти зростання кількості особин популяції.

1. Розмір початкової популяції: *N*=1.На перших 20 ітераціях кількість особин зростає вдвічі. Далі – стала кількість особин..
2. Розмір початкової популяції: *N*=10.На перших 20 ітераціях кількість особин зростає вдвічі. Далі – стала кількість особин.
3. Розмір початкової популяції: *N*=1. Кількість особин зростає з коефіцієнтом 1.005. Період – 3500 ітерацій. Далі – стала кількість особин.
4. Розмір початкової популяції: *N*=10. Кількість особин зростає з коефіцієнтом 1.005. Період – 3500 ітерацій. Далі – стала кількість особин.
5. Розмір початкової популяції: *N*=100. Кількість особин зростає з коефіцієнтом 1.005. Період – 3500 ітерацій. Далі – стала кількість особин.
6. Довжина хромосоми: *l*=385.Розмір початкової популяції: *N*=734. Кількість особин зростає з коефіцієнтом 1.00082. Поріг зростання - розмір популяції 16958. Далі – стала кількість особин.

**Типи відбору:**

- рулетки (RWS)

- турнірний з параметром t=2

- турнірний з параметром t=4

* турнірний з параметром t=12

**Генетичні оператори, що використовувалися в алгоритмі:**

Щільнісна мутація: кожен біт з заданою ймовірністю *pm* змінює значення на протилежне.

Значення *Px* буде надано згодом; в іншому разі підібрати, починаючи з *Px*=1/(10\**l*).

Розглянути випадки:

*pm*= *PX*;

*pm*= *PX+*0.2\**PX*;

*pm*= *PX*–0.2\**PX*;

*pm*= *PX*/2;

*pm*= *PX*/10;

*pm*= *PX*/100;

**Для оцінювання:**

1. Коефіцієнт пристосованості у всіх особин *l*.
2. Для обчислення коефіцієнта пристосованості особин потрібно визначити локуси з нейтральними, патогенними та летальними мутаціями. Розглядати варіанти з *l*>=100. Коефіцієнт пристосованості у всіх особин визначається за наявністю мутацій. За відсутності мутацій (по відношенню до ланцюжка, що складається з усіх 0) або за наявності нейтральних мутацій коефіцієнт пристосованості ***l***. Тут: *l* – довжина ланцюжка.  
   ***Допустимими (нейтральними)*** є мутації в перших 13.5% локусів та ще в наперед визначених 24.5% локусів.   
   ***Патогенними*** (знижують коефіцієнт пристосованості до **(*l*–*k*)** є мутації в наперед визначених 2.32% локусів. Тут: *l* – довжина ланцюжка, *k* – кількість патогенних мутацій.   
   Мутації в інших локусах є ***летальними*** (знижують коефіцієнт пристосованості до ***0.1***)***.***

# **2. Завершення роботи**

*Генетичний алгоритм* зупиняє свою роботу, якщо справджується одне з двох.

* Ідентифіковано збіжність алгоритму: середнє здоров’я популяції не змінюється останні 10 поколінь більше ніж на 0.0001.
* Проведено **20 000** ітерацій (поколінь, породжених ГА).

3. **Аналіз алгоритму.**

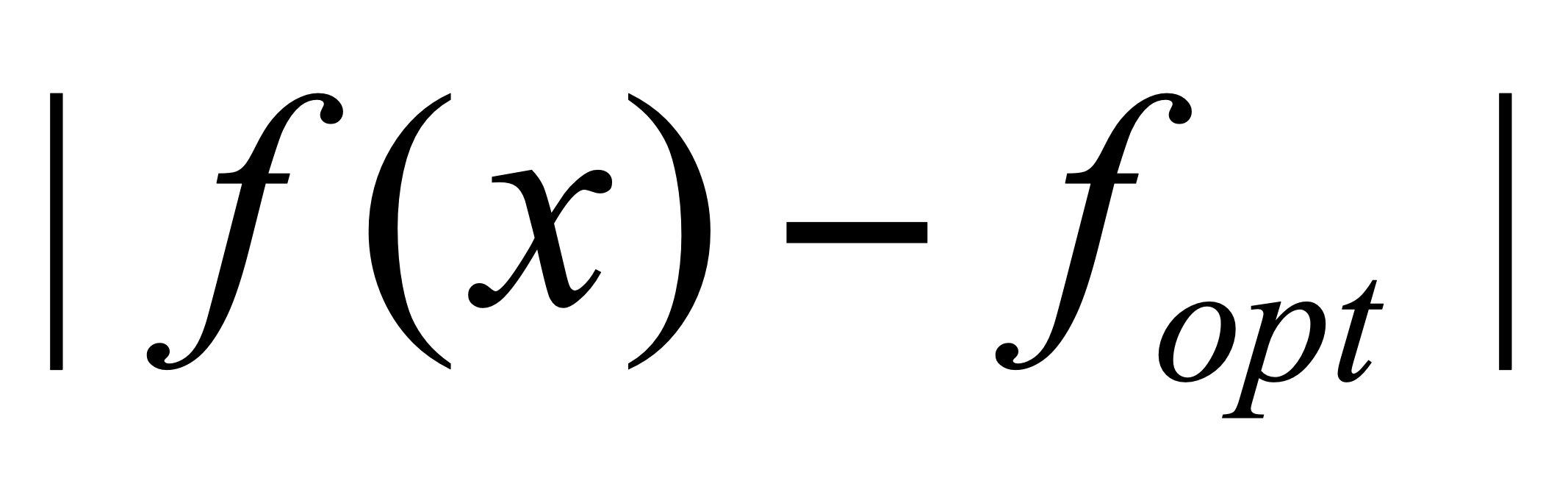
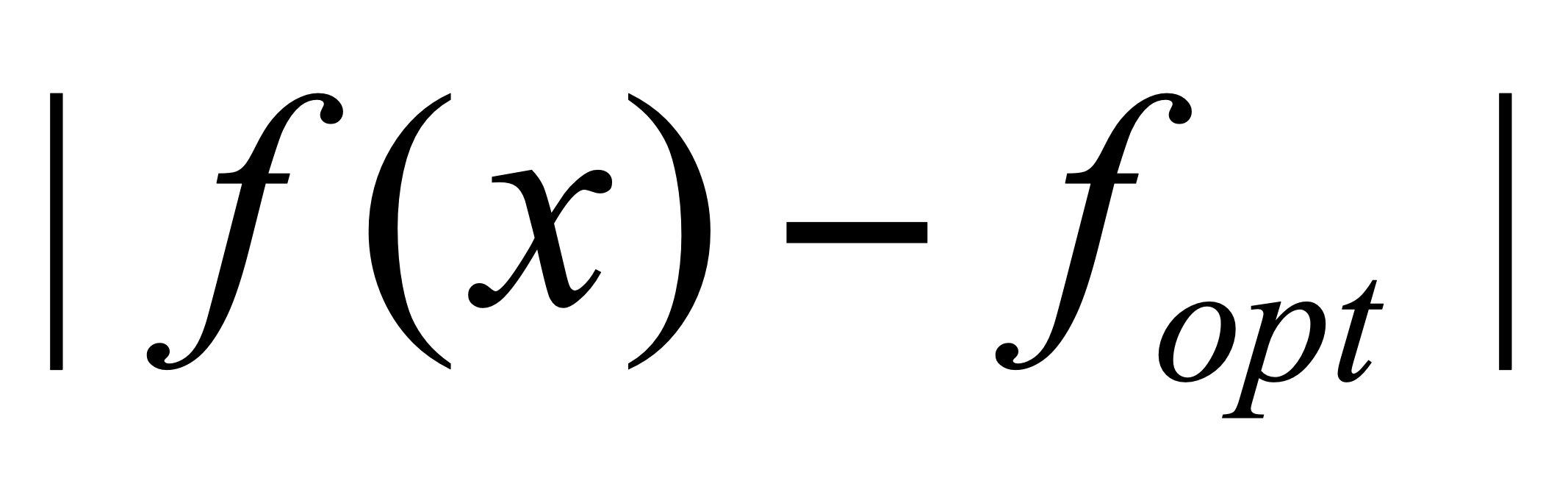
При аналізі стохастичних методів необхідно спиратись не на результати одного прогону, а **генерувати статистику** – усереднені значення багатьох (5) незалежних прогонів одного алгоритму; ефективність оцінюється на основі зібраної статистики.

**Важливо:** *i*-ий прогін кожного алгоритму має одну стартову точку (випадково згенерована початкова популяція однакова для всіх алгоритмів на відповідному прогоні).

Кількість прогонів на одному наборі параметрів, на одній задачі: 5.

**Критерії**

Для **кожного прогону** алгоритму обчислюватимемо такі ***критерії***.

* *NI (number of iterations)* – кількість ітерацій.
* % поліморфних генів (кількість локусів у популяції, по яких значення відрізняється від цільового (залежно від варіанту, кількість «1» або не «А»)) [необхідний для визначення реального значення *pm*]
* середнє значення коефіцієнта пристосованості в популяції;
* найкраще значення коефіцієнта пристосованості в популяції;
* номер ітерації, на якому були отримані задані відсотки різноманітності особин популяції (кількість допустимих непатогенних мутацій): 13.5%+24.5%=38%
* значення відхилення середнього значення коефіцієнта пристосованості від оптимального;
* значення відхилення найкращого знайденого розв’язку від оптимального;
* Зберігати розподіл особин популяції на кожних *h* кроках роботи алгоритму. Цей критерій **ілюструєхмо точковою гістограмою**: **h=300;**

1. Відстані попарні між всіма особинами популяції.
2. Відстані (Геммінга) до цільового ланцюжка (з усіх 0 або А). Іншими словами, слід зберігати відсоток особин популяції без помилок (відстань Геммінга 0), з 1 помилкою (відстань Геммінга 1), з 2 помилками (відстань Геммінга 2), з 3 помилками (відстань Геммінга 3), з 4 помилками (відстань Геммінга 4), і т.д..
3. Відстані до «дикого типу» (обчислюється для кінцевої популяції, «дикий тип» є рядочком, кожен локус якого має найбільш поширене в популяції значення).

По всіх прогонах слід обчислити **усереднені значення** усіх критеріїв (середнє та найкраще по всіх прогонах).

Результати роботи:   
  
Всі результати представлені у репозиторії “<https://github.com/StasSV/Lab_results>"

# Висновок

**Підведемо підсумки завдання №1.**   
Провели дослідження для наступних параметрів (l=100; N=1; коеф. зрост. Популяції = 1.005) з усіма заданими значеннями функції Pm та для (l=100; N=1; коеф. зрост. Популяції = 2) з усіма заданими значеннями функції Pm, “Сардинія”.  
Для всіх прогонів збіжність відбулася до 10000-ї операції.  
Після порівняння турнірного відбору з різними параметрами бачимо, що від збільшення параметру збіжність залежить прямопропорційно (зі зростанням коефіцієнта збіжність зростає).

Також можемо зробити висновок, що при відборі “**Рулеткою**” збіжність відбувається гірше, ніж при “**Турнірному**” відборі.

**Підведемо підсумки завдання №2**. З отриманих результатів можна зробити наступні висновки:   
Під час зростання популяції мутації відбуваються зі зміщенням вправо до моменту стабілізації в лівій частині, після чого пропорція залишається стабільною.