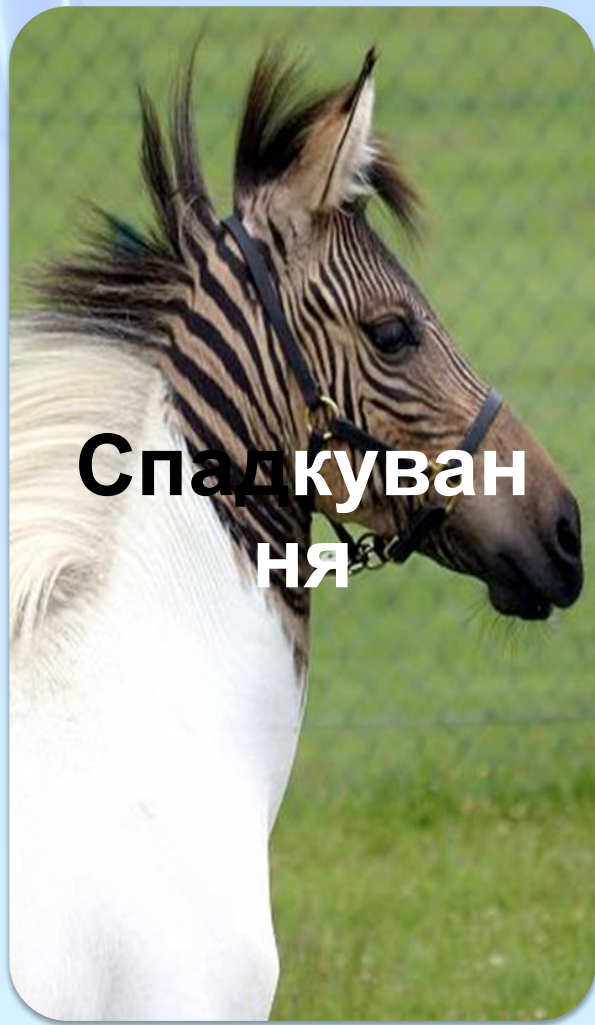


Генетичний

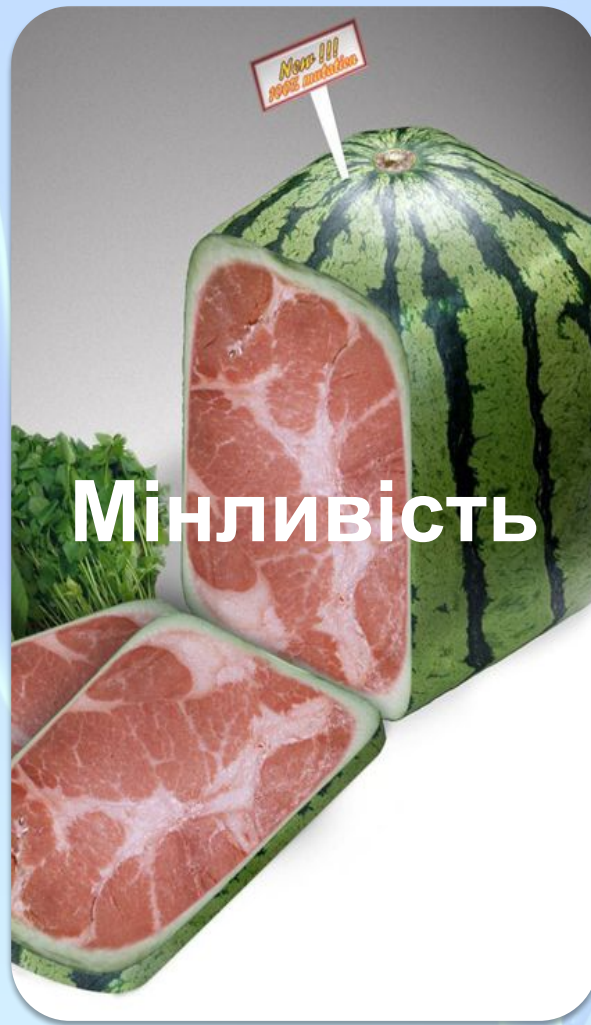


Історія

- Історія еволюційного моделювання або еволюційних обчислень почалася з робіт Дж. Холланда, Л. Фогеля, А.Овена та М.Уолша.
- Усі вони взяли за основу перетворення живих організмів в природі, спростили їх та розробили ряд принципів та моделей еволюційних процесів.
- Згодом еволюційне моделювання перетворилося в теорію, на основі якої виконується пошук оптимальних, або близьких до них розв'язків.



Спадкування



Мінливість



Природний відбір

Теорія еволюції Дарвіна

Принципи розвитку виду

- *Спадковість* (здатність організмів передавати свої ознаки потомству)
- *Мінливість* (забезпечує генетичну різноманітність популяції і має випадковий характер):
 - *комбінаційна* (результат рекомбінації генів в результаті схрещування).
 - *мутаційна* (мутагенез):
 - природна (спонтанна)
 - штучна (індукована).
- Природний відбір (\Rightarrow адаптація і видоутворення)

Основні поняття

- **Ген** - структурна и функціональна одиниця спадковості, що контролює розвиток певної ознаки чи властивості.
- **Хромосома** - сукупність генів, яка характеризує особину.
- **Популяція** – сукупність усіх особин.
- **Локус** – позиція у хромосомі.
- **Алель** – можливі значення гена (сукупності генів, що йдуть поспіль).

Генетичний алгоритм в біології приблизно працює так.

Мама і тато будуть:

Колір очей:

синій

Колір

волосся:

русий

Колір шкіри:

світлий



Колір очей:

синій

Колір волосся:

русий

Колір шкіри:

світлий

В результаті схре

а все виглядатиме так:



Колір очей:

синій

Колір волосся:

чорний

Колір шкіри:

смуглий

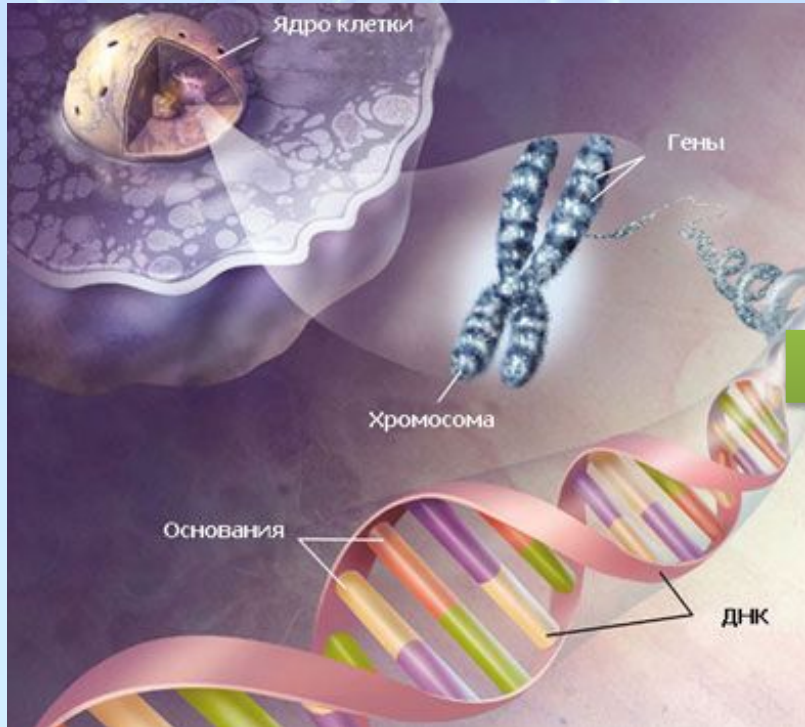
Наочна генетика)))



DEMOTIVAT ORS.T.O

Міждисциплінарний підхід

Біологія



Математика



Генетичний алгоритм

При розв'язанні задачі із застосуванням ЕОМ необхідно :

- **обрати спосіб представлення розв'язку** (необхідно розробити таку структуру, яка дозволить кодувати будь-який можливий розв'язок і проводити його оцінку = спосіб обчислення ЦФ);
- **розробити оператори створення (генерації) особин;**
- **розробити оператори випадкових змін;**
- **визначити закони виживання розв'язків (особин);**

Общая схема алгоритма

Пусть имеем комбинаторную задачу оптимизации:

$$f(x) \rightarrow \max,$$

$$x \in X.$$

Начальная популяция $\{x^1, x^2, \dots, x^k\}$ - набор допустимых решений исходной задачи (особей).

Шаг эволюции:

- выбираем из популяции два решения (родителей),
- скрещиваем их (получаем потомка (-ов)),
- применяем мутацию,
- применяем локальное улучшение,
- добавляем потомка (-ов) в популяцию,
- удаляем из популяции наихудшее (-ие) решение.

Задача про ранець



Задача про ранець



Вхідні дані до задачі

Потрібно скласти ранець в похід так, щоб він мав вагу не більше

15 кг, а його вміст був максимально корисним.

Номер	Назва	Цінність	Вага
1	Спальний мішок	7	2
2	Книжка	2	3
3	Консерва	5	5
4	Ноутбук	1	6
5	Змінне взуття	6	4
6	Сірники	9	2
7	Крупа	8	3

Початкова популяція

(початкові варіанти пакування

ранця)

								
1	0	1	0	0	1	1	29	12
1	0	1	0	1	0	1	26	14
1	1	0	0	1	0	1	23	12
0	0	0	1	1	1	1	26	15
1	1	1	0	1	1	1	37	19

Функція
придатності







Початкова популяція

								
1	0	1	0	0	1	1	29	12
1	0	1	0	1	0	1	26	14
1	1	0	0	1	0	1	23	12
0	0	0	1	1	1	1	26	15

Вибір батьків

								
1	0	1	0	0	1	1	29	12
0	0	0	1	1	1	1	26	15

Схрещування

						
1	0	1	0	0	1	1
0	0	0	1	1	1	1

1	0	1	0			
				1	1	1

1	0	1	0	1	1	1
---	---	---	---	---	---	---

Характеристики

нащадка

								
1	0	1	0	1	1	1	35	14

Оновлення популяції

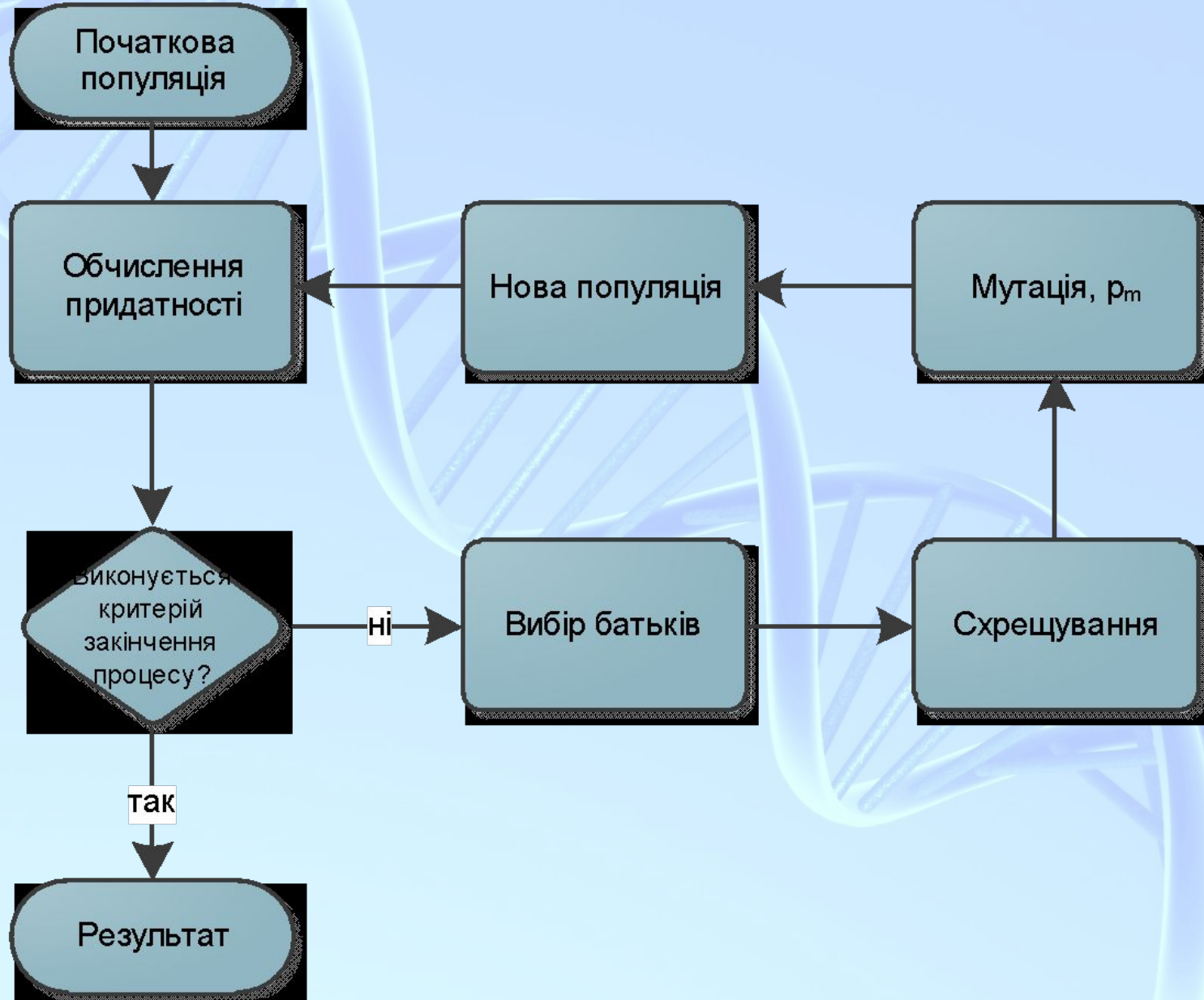
								
1	0	1	0	0	1	1	29	12
1	0	1	0	1	0	1	26	14
1	1	0	0	1	0	1	23	12
0	0	0	1	1	1	1	26	15
1	0	1	0	1	1	1	35	14

Поточний рекордний розв'язок

ЯЗОК



Схема простого генетичного алгоритму



Генерація нащадків

Природне середовище

- **Статеве**

(генетичний матеріал двох батьків використовується при створенні нащадка)

- **Безстатеве**

(клонування, при якому відбуваються різні зміни при передачі інформації від батька до нащадка)

Інтелектуальна штучна система

- **Статеве**

(можна використовувати генетичний матеріал двох, трьох і більше батьків; проводити голосування за батьків тощо)

- **Безстатеве**

(клонування)

Оператори вибору батьків

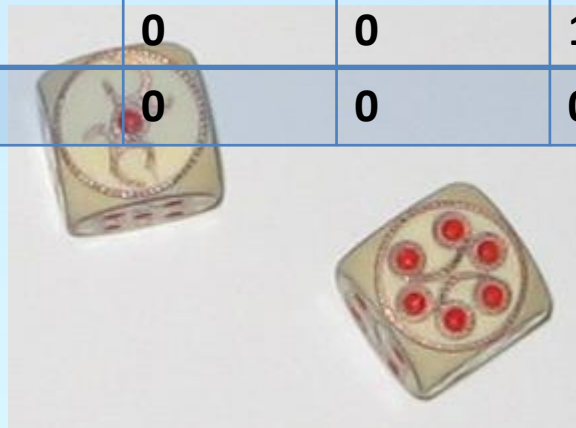
Маємо

0	1	0	0	1	1
1	0	1	1	0	1
1	1	0	1	1	0
1	1	0	1	0	1
1	0	1	1	1	1
1	1	0	0	0	0

Спосіб 1.

Панміксія

0	1	0	0	1	1
1	1	0	0	0	0



Оператори вибору батьківів

Маємо

популяцію

I	0	1	0	0	1	1	10
II	1	0	1	1	0	1	21
II	1	1	0	1	1	0	25
I	1	1	0	1	0	1	15
I	1	0	1	1	1	1	31
II	1	1	0	0	0	0	5



Спосіб 2. Турнірний відбір



Придатність
10



Придатність
21



Придатність
25



Придатність
15



Придатність
31



Придатність
5

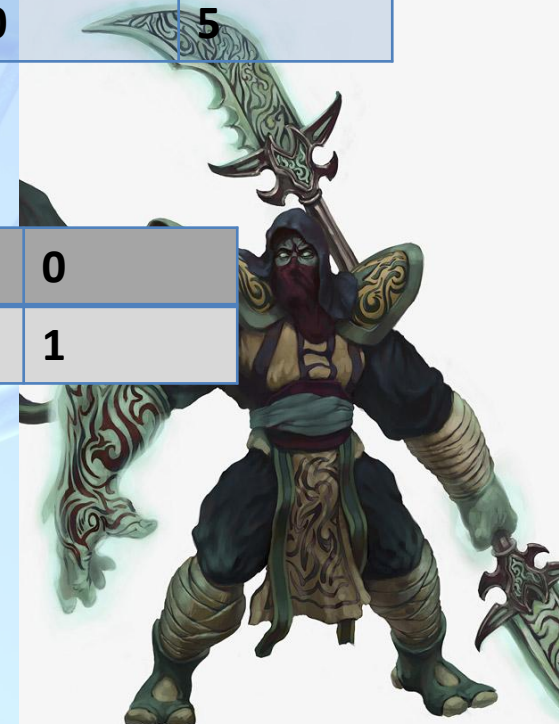
Способи вибору батьків

Маємо
популяцію

I	0	1	0	0	1	1	10
II	1	0	1	1	0	1	21
II	1	1	0	1	1	0	25
I	1	1	0	1	0	1	15
I	1	0	1	1	1	1	31
II	1	1	0	0	0	0	5

посіб 2. Турнірний
дбір

1	1	0	1	1	1	0
1	0	1	1	1	1	1



Оператори вибору батьків

- *Панміксія*
- *Турнір*
- *Інбридинг* (першого батька відбирають випадковим чином, а другий батько є найближчим до першого членом популяції)
- *Аутбридинг* (шлюбні пари формують з максимально віддалених особин)
- *Селекція*¹ (особини, значення пристосованості яких не менше порогової величини)
- *Метод рулетки*
 - випадковий + випадковий, де

$$p_i = \frac{f_i}{\sum_{j=1}^n f_j} \quad - \text{ймовірність вибору } i\text{-ї особини}$$

- випадковий з топ-множини + випадковий
- кращий + випадковий

¹ Природна селекція – процес, при якому хромосоми з більш сильними ознаками (кращою ЦФ) мають більшу можливість для репродукції, ніж слабі

Оператори вибору батьків

Інбридинг

Приводить до концентрації пошуку в локальних ділянках, що фактично призводить до розбиття популяції на окремі локальні групи навколо підозрілих на екстремум ділянок МДР

генотипний¹

фенотипний²

Аутбридинг

Спрямований на попередження збіжності алгоритму до вже знайдених розв'язків, змушуючи алгоритм переглядати нові, недосліджені області

генотипний¹

фенотипний²

¹ в якості відстані береться різниця значень цільової функції

² як відстань береться відстань Хемінга

Відстань Хемінга

*Хемінгова відстань між хромосомою **1010001**
і хромосомами популяції*

Хромосоми популяції	Кількість локусів, що відрізняються
1000000	2
1010101	1
1111111	4
1100001	2
0110011	3

Оператори створення (генерації) особин

Призначення: створювати нащадків на основі схрещування батьків

Назви:

- оператор схрещування
- оператор кросинговеру
- оператор кросоверу
- оператор рекомбінації

Структура операторів в основному визначає **ефективність ГА**

Оператор рекомбінації (схрещування)

Бінарна рекомбінація (кросинговер)

Два

батька:

0	1	0	0	1	1
1	1	0	0	0	0

Одноточковий кросинговер ↑ :

0	1	0	0	1	1
---	---	---	---	---	---

1	1	0	0	0	0
---	---	---	---	---	---

Розрізняють такі види

кросинговеру:

- ✓ Одноточковий кросинговер.
- ✓ Двоточковий кросинговер.
- ✓ Багатоточковий кросинговер.

Оператор рекомбінації

(схрещування)

- Бінарна рекомбінація
(кросинговер)

Два
батьки:

0	1	0	0	1	1
1	1	0	0	0	0



Двоточковий
кросинговер:

Два
нащадки:

0	1	0	0	0	1
1	1	0	0	1	0

Точки кросинговеру

можуть:

- ✓ бути постійними;
- ✓ обиратися випадково.

Оператор рекомбінації (схрещування)

Дискретна рекомбінація

Два

батька:

123	75	34	86
78	21	54	22

Класична дискретна

рекомбінація

1	2	1	1
2	1	2	2

Отримані

діти:

123	21	34	86
78	75	54	22

Оператор рекомбінації (схрещування)

Дискретная рекомбинация

Родитель 1 (Репликация) Родитель 2

1.77	↔	8.84
-4.80	↔	5.49
-3.49	↔	1.49
8.76	↔	6.03
-1.49	↔	-0.57

Репликация. Оператор наиболее близок к природному явлению, которое в биологии имеет название «Репликация ДНК».

Репликация является важнейшим генетическим оператором, который генерирует новые гены, при этом передает признаки родительских хромосом.

Оператор рекомбинации (Репликация)

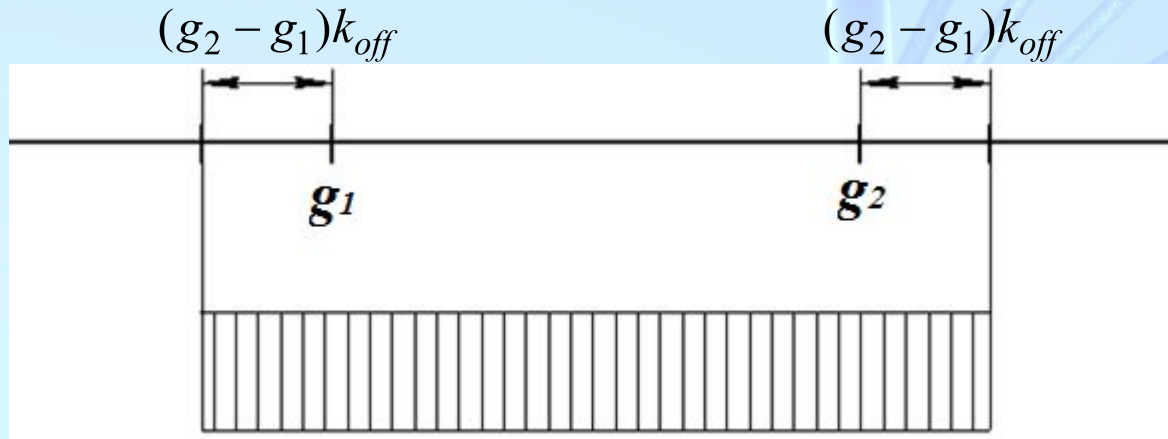
Родитель1 Родитель2

1.77	↔	8.84
-4.80	↔	5.49
-3.49	↔	1.49
8.76	↔	6.03
-1.49	↔	-0.57

Пусть g_1, g_2 - значения генов родителей 1 и 2 соответственно ($g_1 < g_2$);
 k_{off} - коэффициент смещения (расширения) границ

интервала . $g_1 - (g_2 - g_1)k_{off} \div g_2 + (g_2 - g_1)k_{off}$

Ген потомка - случайное число из интервала:



Оператор рекомбинации (Репликация)

Родитель1 Родитель2

П

1.77	↔	8.84	4.57
-4.80	↔	5.49	6.34
-3.49	↔	1.49	-1.01
8.76	↔	6.03	5.84
-1.49	↔	-0.57	-1.20

Оператор рекомбинации

Універсальний оператор кросинговеру

Популярний при розв'язанні задач теорії розкладів. Замість використання точок розрізу (точок кросинговеру) визначають двійкову маску, довжина якої дорівнює довжині хромосоми. Отримання нащадків виконується на основі бінарної операції «додавання по модулю 2» («исключающее ИЛИ») над генами батьків і маски: $0+0=0$; $0+1=1$; $1+0=1$; $1+1=0$

Батько 1	0	1	1	0	0	1	0	1
Батько 2	0	1	0	1	1	1	1	1
МАСКА	0	1	1	0	0	0	1	1
Нащадок 1	0	0	0	0	0	1	1	0
Нащадок 2	0	0	1	1	1	1	0	0

Маска зазвичай формується випадковим чином

Мутація

- З'являються постійно в ході процесів, які відбуваються в живій клітині
- Виникають мимовільно протягом усього життя організму (з частотою одного разу на 10^{10} клітинних генерацій).
- Є матеріалом для природного відбору.

Служить для:

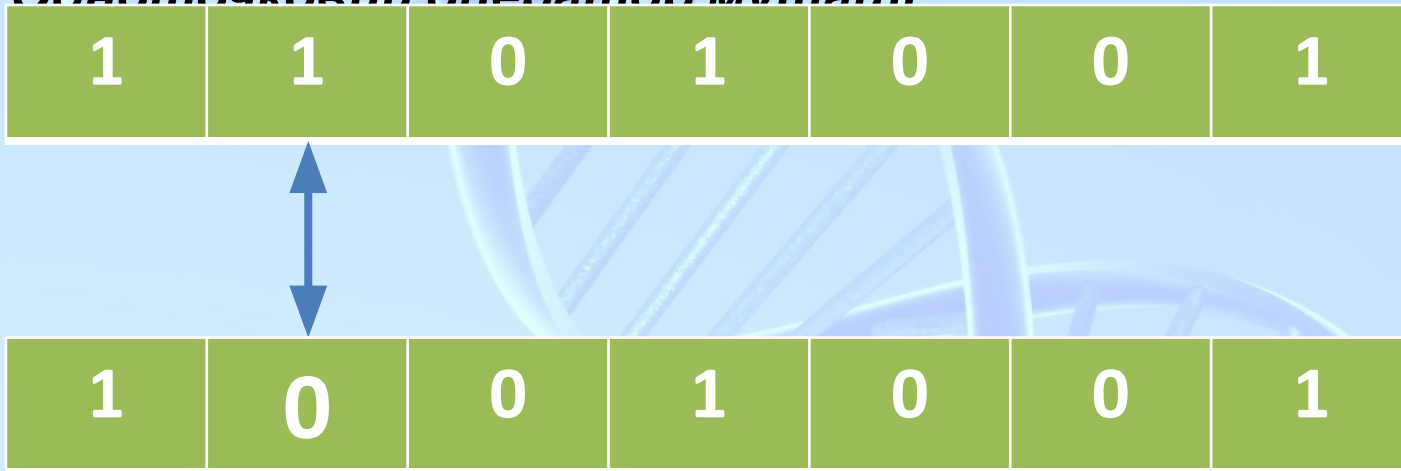
- “Вибивання” популяції з локального екстремуму
- Запобігання передчасної збіжності
- Прискорення пошуку оптимального розв'язку



Мутація

Випадок використання бітового кодування генів та хромосом

Одноточковий оператор мутації



Мутація

Випадок використання бітового кодування генів та хромосом

Двохточковий оператор мутації



Мутація

Позиційний оператор мутації

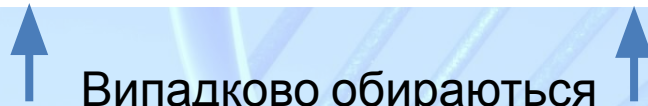


↑
Випадково обираються
дві точки мутації
↑

Ген, що відповідає другій точці мутації, розміщується в позицію перед геном, якій відповідає першій точці мутації



Інверсія



Випадково обираються
дві точки інверсії

Гени, що розміщені між цими точками,
інвертуються



Мутація

*Дискретний
випадок*

20	33	45	16	37	81	52
----	----	----	----	----	----	----

Спосіб

Нова змінна = стара змінна $\pm \alpha \cdot \delta$

де α - деяка константа δ - випадкове число.

Наприклад, $\alpha = 5, \delta = 0,4$

-2	+2	+2	-2	+2	-2	-2
----	----	----	----	----	----	----

Результат:

18	35	47	14	39	79	50
----	----	----	----	----	----	----



Мутація

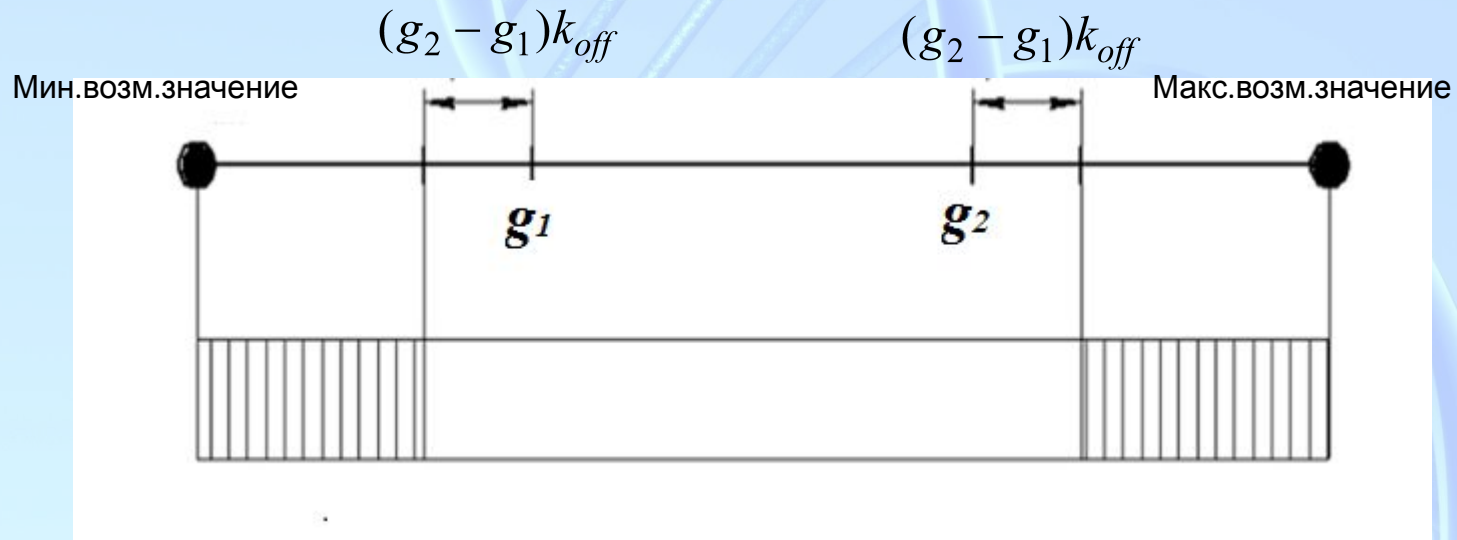
Дискретный случай

Способ 2

Пусть g_1, g_2 - значения генов родителей 1 и 2 соответственно ($g_1 \leq g_2$);
 k_{off} - коэффициент смещения границ интервала.

Ген потомка после мутации: случайное число **вне интервала**:

$$g_1 - (g_2 - g_1)k_{off} \div g_2 + (g_2 - g_1)k_{off}$$



Реанімація



Відбір особин в нову популяцію

(задача на мінімум)

Батьк

0	1	0	0	1	1	10
1	0	1	1	0	1	21
1	1	0	1	1	0	25
1	1	0	1	0	1	15
1	0	1	1	1	1	31
1	1	0	0	0	0	5

Нащадк

0	1	0	1	0	1	16
1	1	0	1	1	1	29

Відбір особин в нову популяцію

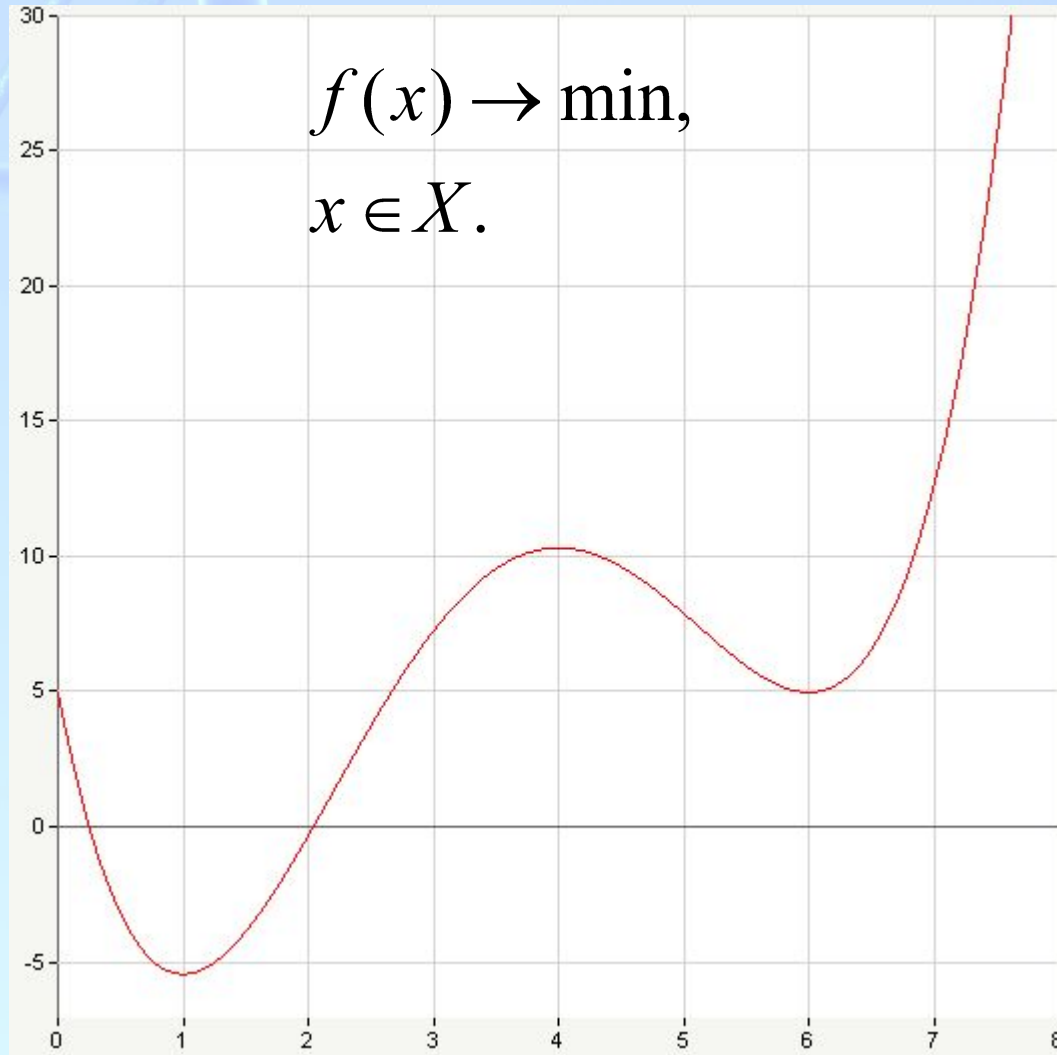
(задача на мінімум)

Нова популяція	1	0	0	1	1	10
1	0	1	1	0	1	21
1	1	0	1	1	0	25
1	1	0	1	0	1	15
0	1	0	1	0	1	16
1	1	0	0	0	0	5

Способи відбору особин в нову популяцію

- *Елітарний відбір* (серед всіх елементів популяції і нащадків вибираються N найкращих (найбільш придатних))
- *Відбір витісненням* (вибір особини в нову популяцію залежить не тільки від величини її придатності, але й від того, чи є вже в популяції особина з аналогічним хромосомним набором)

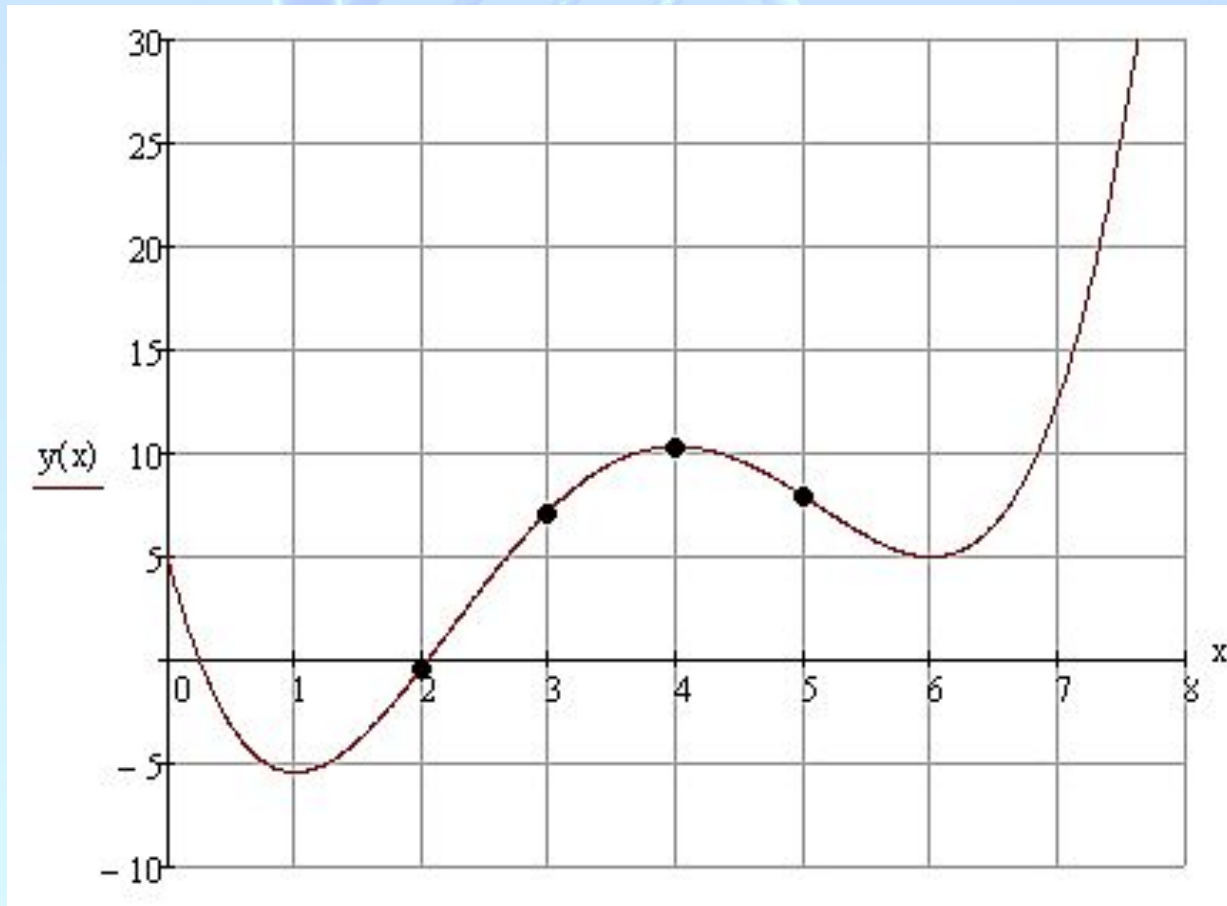
Простий приклад застосування генетичного алгоритму



Знайти глобальний мінімум функції на МДР:

$$f(x) = 5 - 24x + 17x^2 - \frac{11}{3}x^3 + \frac{1}{4}x^4 \rightarrow \min, x \in X = [0;7].$$

Формування початкової популяції: оберемо випадковим чином кілька чисел на відрізку $x = [0;7]$: $\{2,3,5,4\}$.



Тепер приступимо до процесу *розмноження (рекомбінація)* : спробуємо на основі вихідної популяції створити нову.

№	Особина		
	Ціле число	Двійкове число	Пристосованість
1	2	010	-0,33
2	3	011	7,25
3	5	101	7,92
4	4	100	10,33

Процес створення першого покоління потомків має вигляд:

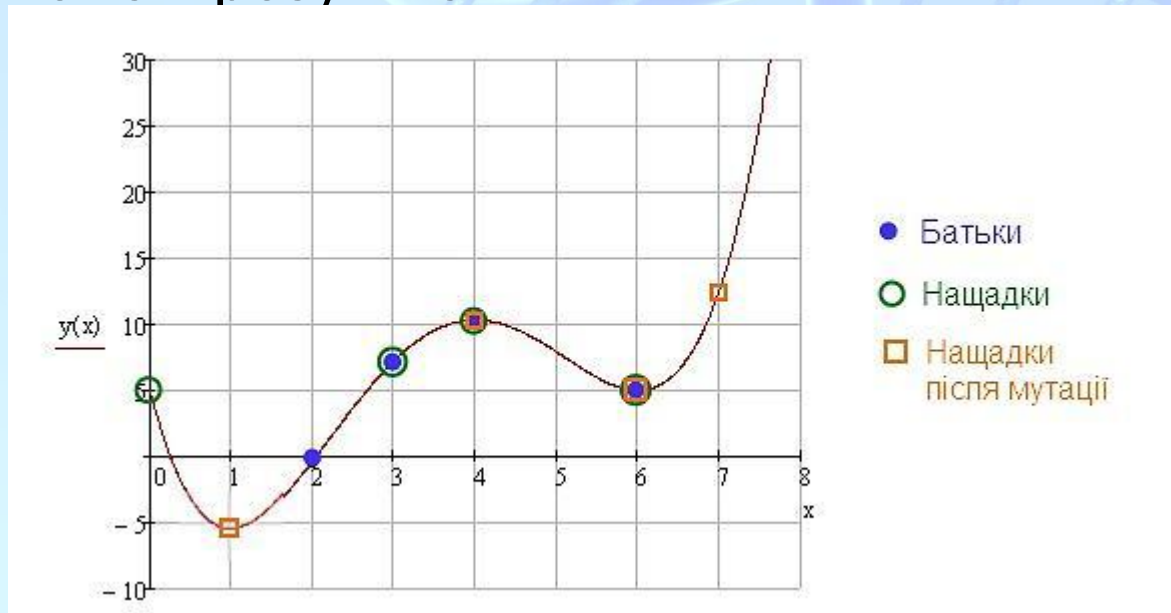
№ нащадка	Особини-батьки		Точка кросинговеру	Особини-нащадки
	010	100		
1	010	100	1	000
2				110
3	101	010	2	100
4				011

Хромосома є зв'язком генів (алелів), розташованих на певній фізичній позиції (локусі) хромосоми та називається **геном**.

Наступним кроком у роботі генетичного алгоритму є мутації
Нехай імовірність мутації дорівнює 0,3.

№	Особини-нащадки	Випадков е число	Ген для мутації	Нащадок після мутації	Пристосованість нащадка до мутації	Пристосованість нащадка після мутації
1	000	0,1	3	001	5	-5,42
2	110	0,6	-	110	5	5
3	100	0,5	-	100	10,33	10,33
4	011	0,2	1	111	7,25	12,58

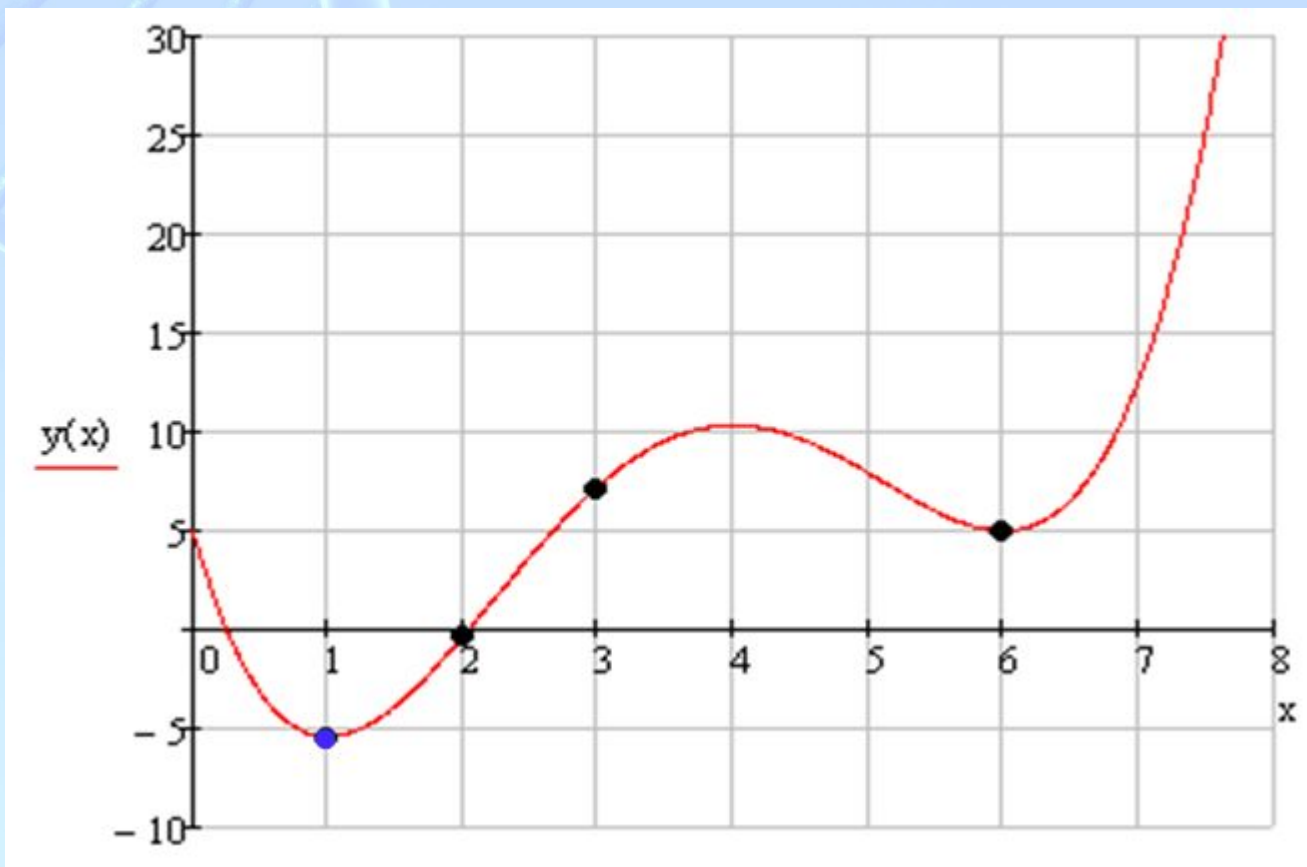
Тепер маємо такі результати:



Тепер з чотирьох особин-батьків і чотирьох отриманих особин-нащадків необхідно сформуванати нову популяцію.

№	Особини	Приста́сованість	Нова популяція	Приста́сованість особин в новій популяції
1	010	-0,33	001	-5,42
2	011	7,25		
3	101	7,92	010	-0,33
4	100	10,33		
5	001	-5,42	110	5
6	110	5		
7	100	10,33	011	7,25
8	111	12,58		

У результаті отримаємо нове покоління, яке представлене на рисунку.



Мінімумом цільової функції є значення -5,42, яке відповідає аргументу $x = 1$.

Переваги і недоліки генетичного алгоритму

- проблематично знайти точний глобальний оптимум;
- генетичний алгоритм неефективно застосовувати у разі оптимізації функції, яка вимагає великого часу на обчислення;
- генетичний алгоритм непросто змоделювати для знаходження всіх розв'язків задачі;
- не для всіх задач вдається знайти оптимальне кодування параметрів;
- важко застосувати для *ізолюваних* функцій.

Ізолюваність («пошук голки в копиці сіна») - проблема для будь-якого методу оптимізації, оскільки функція не надає ніякої інформації, не підказує, в якій області шукати максимум,). Лише випадкове попадання особини в глобальний екстремум може вирішити задачу

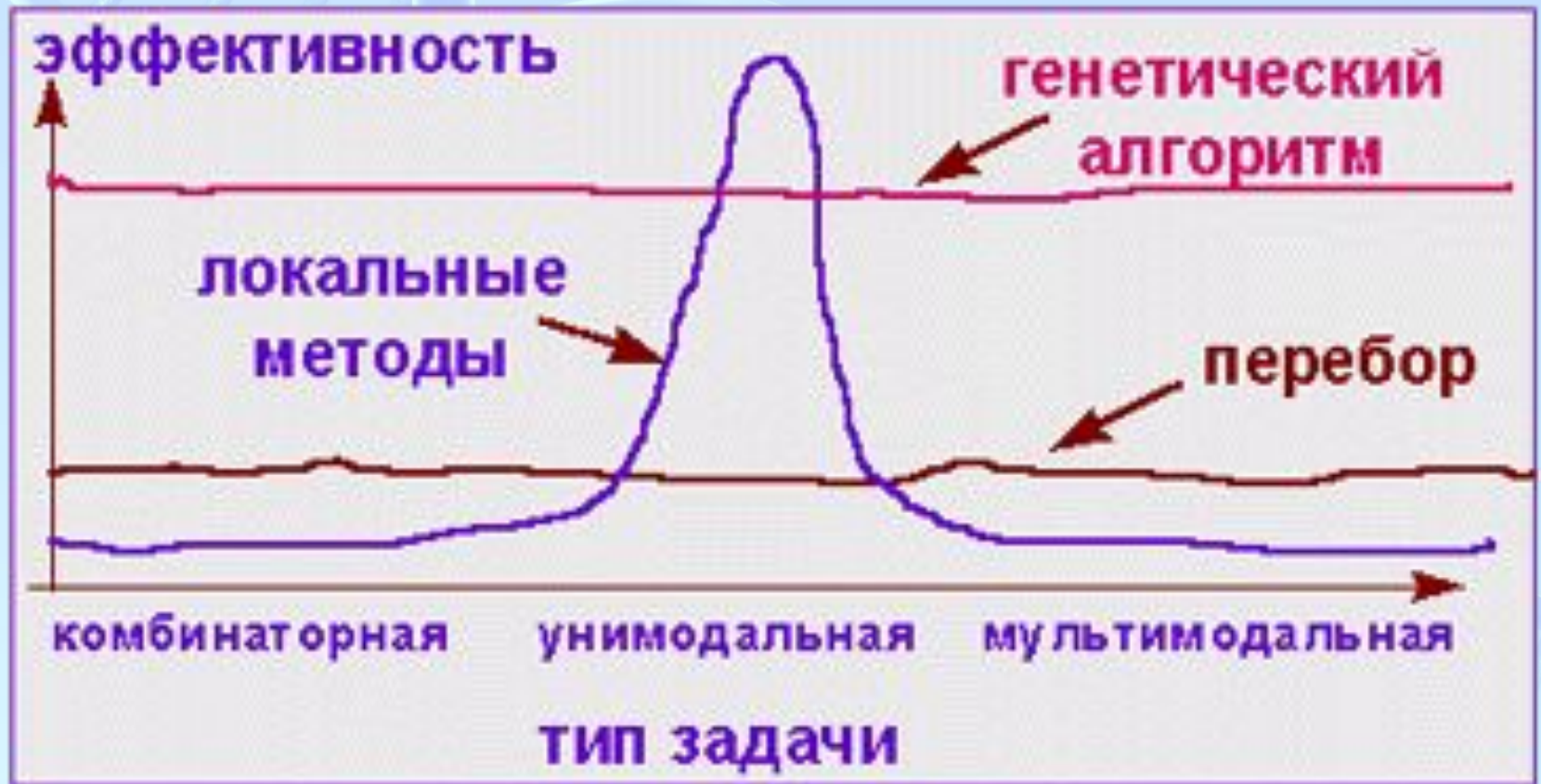


Переваги і недоліки генетичного алгоритму

- + стійкий до потрапляння в локальні оптимуми;
- + придатний для вирішення великомасштабних задач оптимізації;
- + може бути використаний для широкого класу задач;
- + простий в реалізації;
- + може бути використані в задачах з мінливих середовищем.

Ефективність алгоритму сильно залежить від обраної розробником структури, а значить і від його компетентності в даному питанні!!!

Про ефективність ГА



Приклади застосування

- Проектування автомобіля (**Designing the Car**)
 - Ролік <http://boxcar2d.com/>
 - Опис алгоритму
<http://boxcar2d.com/about.html>
- Генерація дизайнерських ідей за допомогою ГА
(<http://habrahabr.ru/company/luxoft/blog/150966/#habracut>)

ГЕНЕТИЧЕСКИЙ БУМ

- Благодаря распространению компьютеров применять "цифровой дарвинизм" стали не только математики.
- В 1994 году Эндрю Кин из университета Саутхемптона использовал генетический алгоритм в дизайне космических кораблей. Взяв за основу модель опоры космической станции, спроектированной в NASA, Кин перевел ее параметры в бинарные строчки-"хромосомы", сделал приличное количество копий и запустил кибер-секс на 11 компьютерах. . После смены 15 поколений, включавших 4.500 вариантов дизайна, получилась модель, превосходящая по тестам тот вариант, что разработали люди.
- В NASA сделали соответствующие выводы: аналогичный генетический алгоритм был использован при разработке антенны для серии микроспутников для исследования магнитосферы Земли.
- Помимо успехов на тестах, разработчиков в обоих случаях поразили органичные формы выращенных конструкций: космическая опора Кина оказалась похожей на берцовую кость, а антенна для спутников NASA - на рог оленя.
- Глава компании Imagination Engines Стивен Талер утверждает, что его программа Creativity Machine уже давным-давно изобрела за него кучу вещей, начиная от зубной щетки Oral-B CrossAction и кончая красивыми названиями типа Synaptrix.
- <http://fuga.ru/articles/2004/03/genetic-pro.htm>