

Генетичні алгоритми

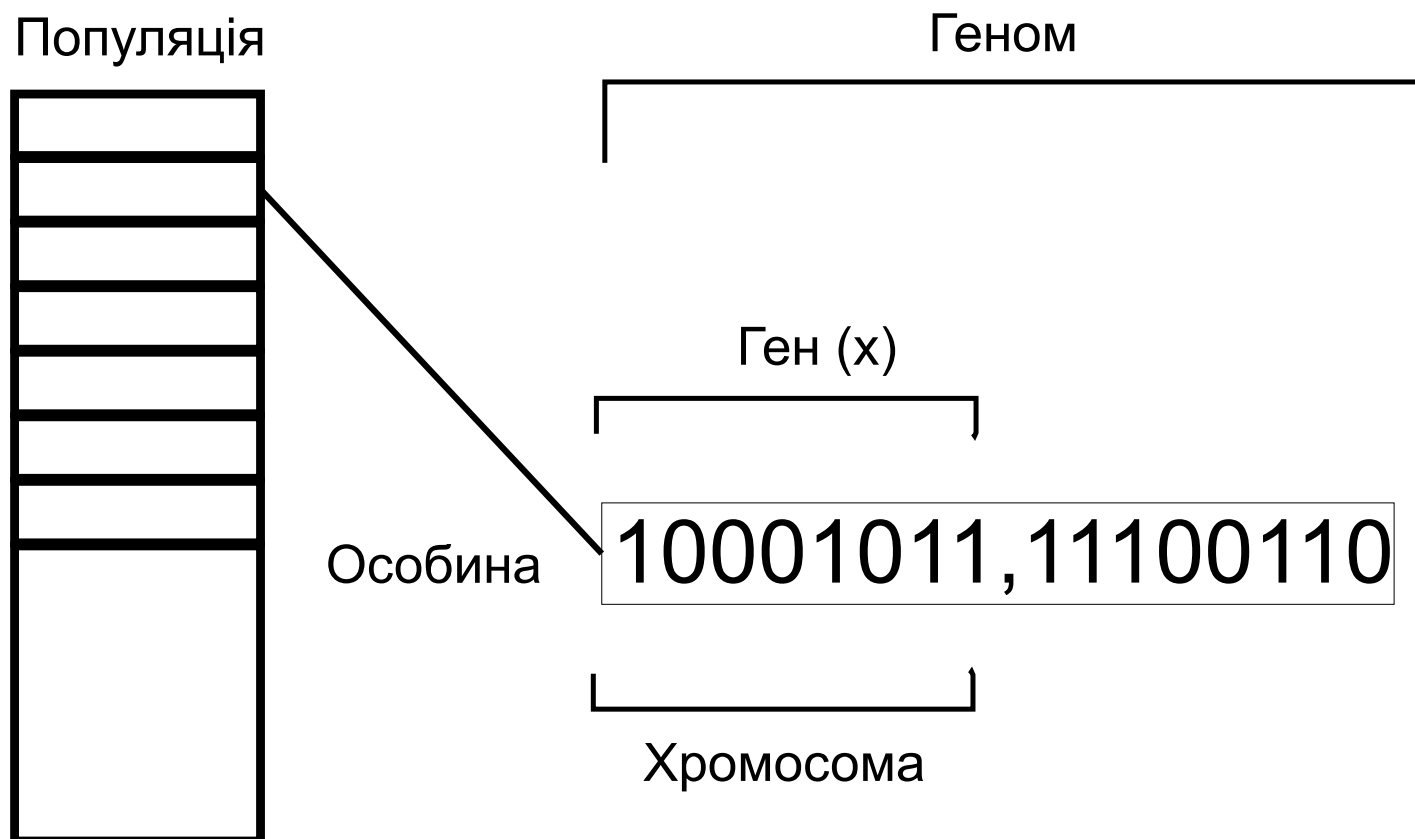
- Універсальні алгоритми оптимізації. Майже незалежні від задачі.
- Добре працюють в складному і зашумленому просторі пошуку.
- Використовують принципи біологічної еволюції
- Базуються на популяції – множині пробних розв'язків (особин)
- Особини взаємодіють між собою, своєрідно обмінюючись інформацією про “стан справ”

Недоліки

- Стохастичний метод
- Слабке теоретичне обґрунтування в більшості реальних випадках
- Повільність. Складно досягнути високої точності.

- Функція для оптимізації: $F(x, y)$.
- **Особина** — пробний розв'язок.
- Загальний вигляд пробного розв'язку: (x_i, y_i) .
- **Фенотип** — пара (x, y) яка задає представлення особини в просторі пошуку
- Генетичне представлення особини - **генотип**. Існує багато варіантів:
 - Бінарний (двійковий) рядок — 10010101. Природне застосування - кодування чисел
 - Рядок символів — *ATGCCGGTAA*. Зручне для деяких задач, наприклад роботи з текстом, дослідження ДНК
 - Дерево. Аналітичні обчислення, аналіз діаграм Фейнмана.
- **Ген** — параметр розв'язку. Для наведеного вище прикладу найбільш природній вибір генів — це x і y .
- **Хромосома** — заповнене представлення гена.
- **Алелі** — можливі варіанти гена.
- **Пристосованість** (fitness) особини i з фенотипом (x_i, y_i) це значення функції $F(x_i, y_i)$. Якість даної особини/розв'язку.

Пошук мінімуму функції $F(x, y)$, $0 < x, y < 17$:



Генотип: $(10001011, 11100110)b \equiv (139, 230)d$

Фенотип: $(139/2^8, 230/2^8) \equiv (0, 54296875; 0, 8984375)$

Пристосованість: $F(x, y) = F(139/2^8, 230/2^8)$

Генетичні оператори — мутація

З певною ймовірністю довільно міняємо один або більше бітів геному:

10011011001



10011001001

Реалізує випадковий пошук — створює нову інформацію. Якщо використати кодування типу коду Грея¹, то може здійснювати локальний пошук.

0	1	2	3	4	5	6	7
000	001	011	010	110	111	101	100

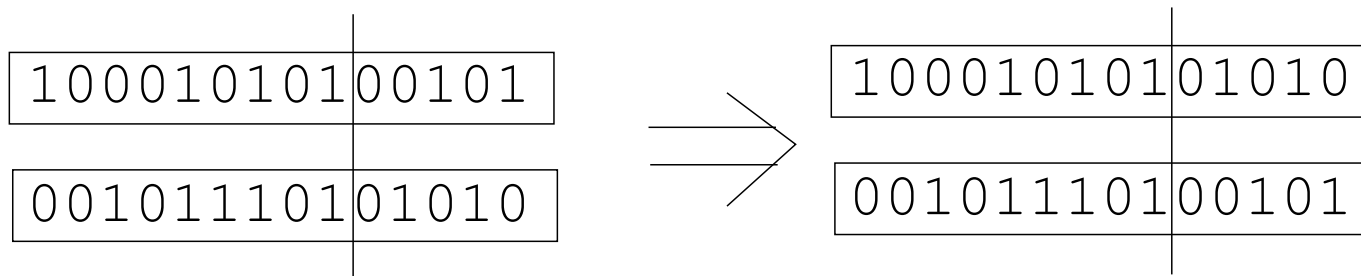
¹Зміна будь-якого біта змінює десяткове число на одиницю

Генетичні оператори — кросингвер

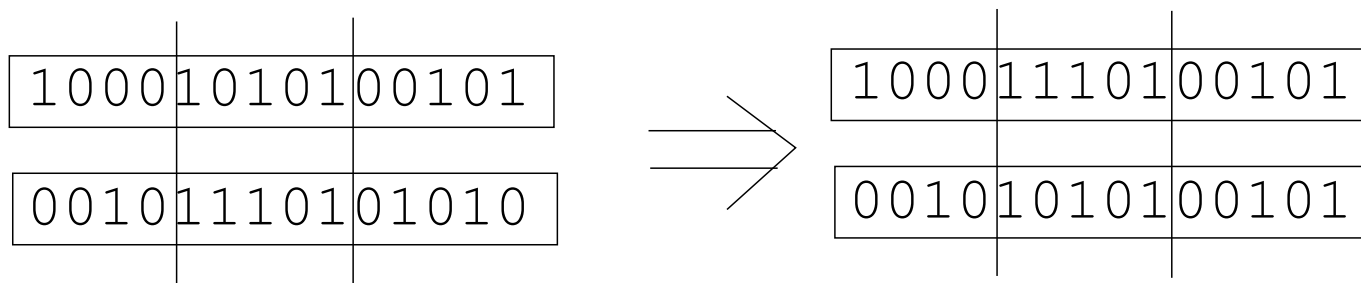
Кросингвер застосовується до пари особин.

Найпростіший випадок:

- Випадковим чином вибираємо точку кросингвера
- Міняємо місцями біти обох рядків після цієї точки:



Двохточковий кросингвер:



Загальна схема

Створюємо популяцію, обчислюємо значення пристосованості її членів

Повторюємо, поки не виконана умова завершення²

- Вибираємо особини для спарювання — кросинговера
- Рекомбінуємо їх — здійснюємо кросинговер
- Застосовуємо мутації
- Обчислюємо значення пристосованості
- Повертаємо отримані особини в популяцію

²досягнуто бажану точність, перевищено задану кількість ітерацій, тощо.

Простий генетичний алгоритм

Властивості

- Двійкове кодування.
- Одноточковий кросинговер
- Репродукція, пропорційна до пристосованості (FPR^3).

Агоритм

Випадково заповнюємо початкову популяцію

Повторюємо, поки не виконано умову завершення:

³Fitness proportionate reproduction

- Створення репродуктивної групи (mating pool). Кожна особина вибирається з ймовірністю, пропорційною до її відносної пристосованості:

$$P(l) = \frac{F(l)}{\sum_{l'} F(l')}.$$

Краще пристосовані особини мають більше шансів бути вибраними.

Особина може бути вибрана декілька раз.

- З ймовірністю p_c до пари особин застосовується кросинговер у випадковій точці.
- До всіх нащадків застосовуються мутації – кожен біт змінюється з ймовірністю p_m .
- Заміщаємо цілу вихідну популяцію нащадками.

Приклад простого ГА – селекція:

No	Особина	Пристосованість	Ймов. бути вибраним	Репр. група
1	10000	81	0.51	10000
2	00101	10	0.06	01011
3	01011	31	0.20	10000
4	01100	36	0.23	01100
Σ		158	1.00	

Приклад простого ГА: $p_c = 0.5$; $p_m = 0.05$

Репр. група	Прист.	Після кросинг.	Прист.	Після мутацій	Прист.
10000	81	10000	81	10001	82
01011	31	01011	31	01011	31
10000	81	10100	90	10100	90
01100	36	01000	27	01000	27
Σ	229		229		230

Критика простого ГА

- Схильний до передчасної збіжності – колапсу популяції до локального мінімуму.
- Не здатен підтримувати помітну різноманітність.
- Популяція досить швидко стає однорідною

Шими (Schemata)

- Шима (schema) над алфавітом A – рядок символів, що належать алфавіту A доповнений символом довільного значення '*'.

Приклад: двійковий алфавіт $A = \{0, 1\}$

- Приклад шими: $11 * * 0$. Їй відповідають чотири рядки:

$\{ 11000, 11010, 11100, 11110 \}$.

- Шима – різновид шаблону.
- ГА насправді насправді працює з *шими* а не з рядками.

Характеристики шим

- Порядок шими $o(S)$ – це кількість заданих позицій.

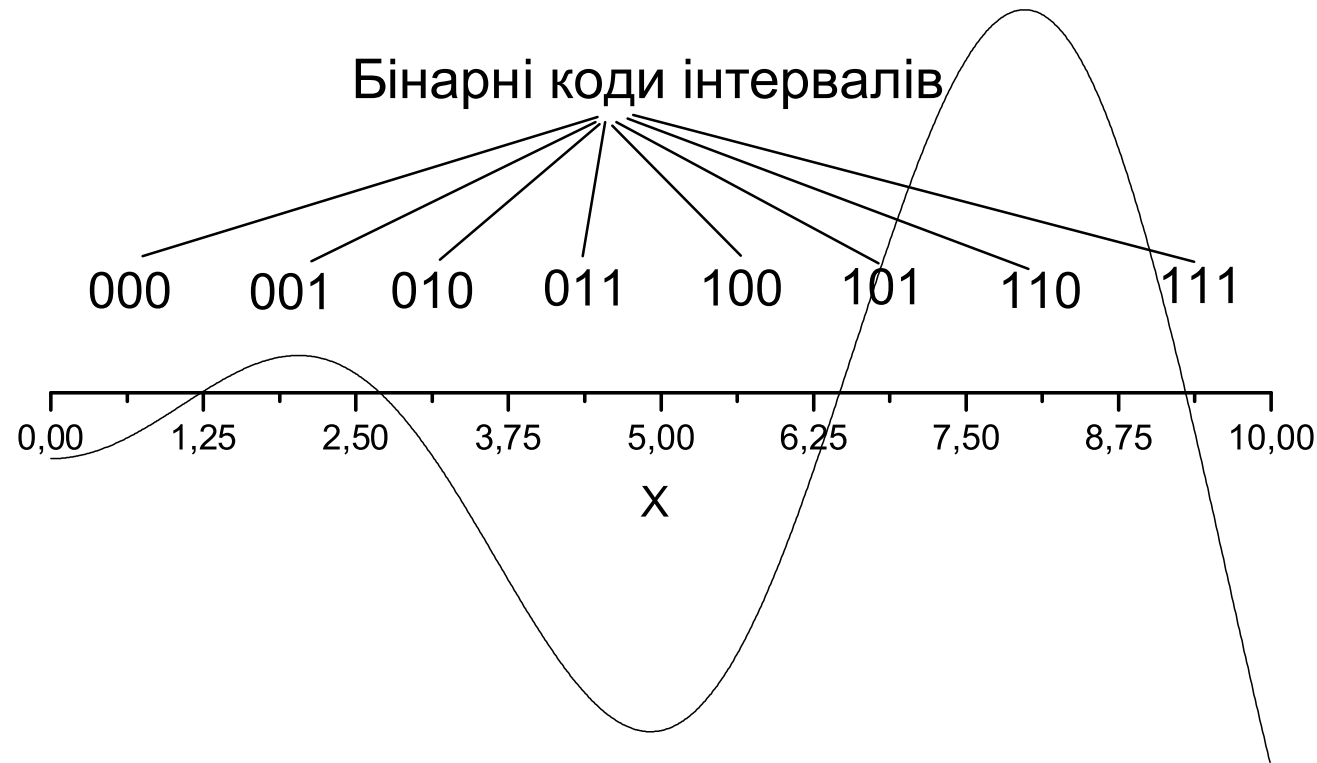
Приклад: $o(11 **0) = 3$; $o(****1) = 1$

- Характена довжина (defining length) $\delta(S)$ – кількість точок для кросинговеру між найдальшими фіксованими точками шими.

Приклад: $\delta(*1**0) = 4$; $\delta(****1) = 0$

Геометрична інтерпретація

$F(x) = 10 + x \sin(x)$, задана на інтервалі $[0, 10]$

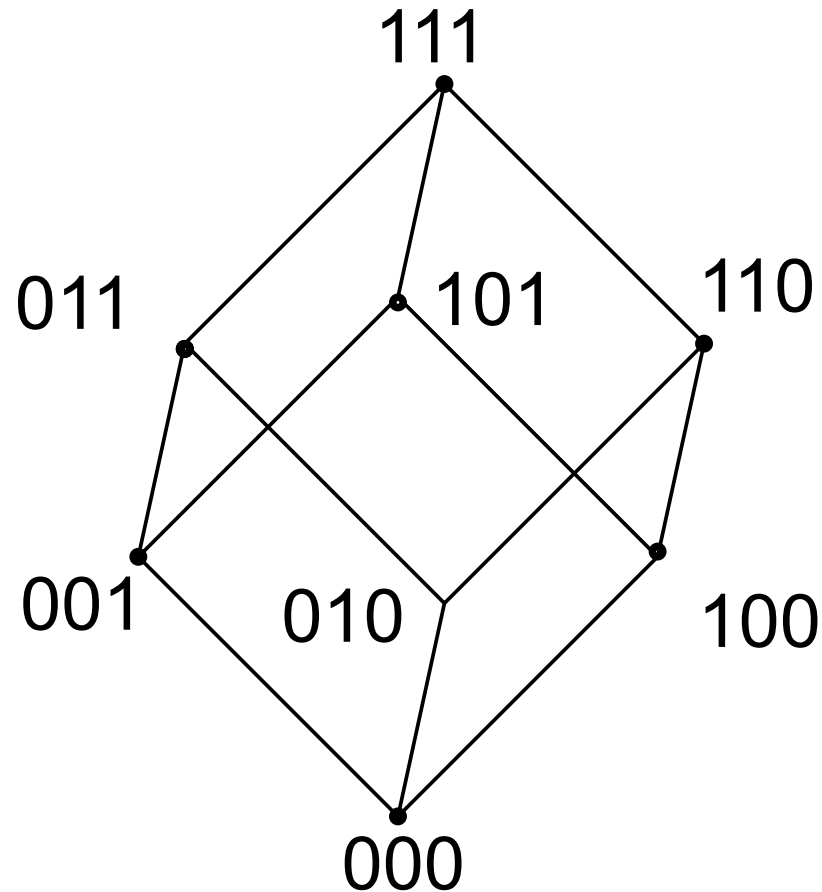


Кількість інтервалів: $2^3 = 8$; їх довжина: $l = 10/8 = 1.25$

Двійковий рядок кодує номер інтервалу.

Геометрична інтерпретація 2

Простір пошуку – множина всіх двійкових рядків з довжиною 3



Популяція здійснює блукання по вершинах цього куба (згідно вибраних правил), досліджуючи набір інтервалів.

Геометрична інтерпретація 3

- Перебір всіх можливостей неефективний, і в більшості задач неможливий.
- Але ГА можуть вибирати гіперплощини⁴ з високою пристосованістю.
- Цим гіперплощинам відповідають шими.
- Біологічна аналогія: геном містить як важливі, консервативні ділянки, мутація яких фатальна⁵ – різко знижує пристосованість організму, так і маловажливі – зміни в яких мало впливають на життєздатність організму.

⁴В просторі пошуку.

⁵Відомо кілька прикладів висококонсервативних генів, які мало відрізняються у людини та бактерій – найбільш філогенетично віддалених організмів.

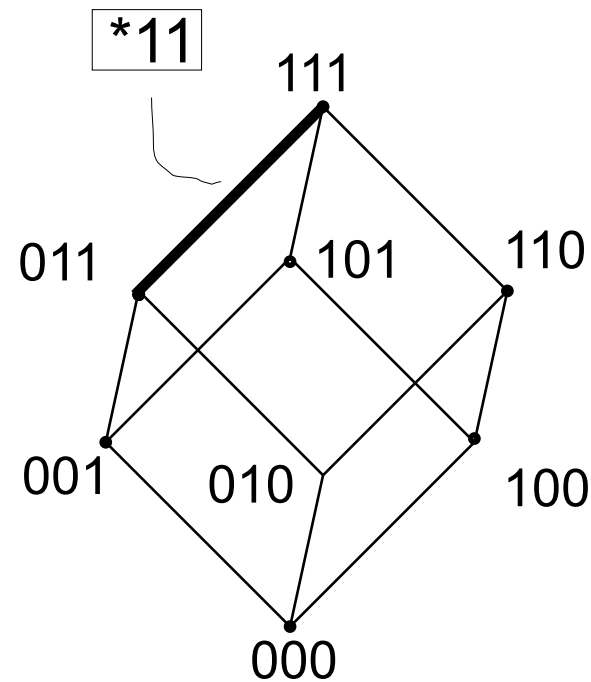
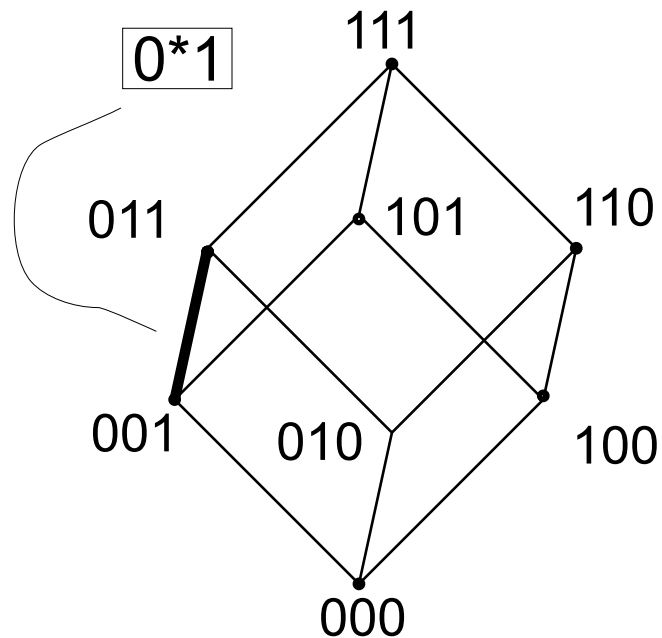
Геометрична інтерпретація 4

Шими 3-го порядку – вершини куба.

Шими 2-го порядку:

$(00*, 01*, 10*, 11*, 0*0, 0*1, 1*0, 1*1, *00, *01, *10, *11)$

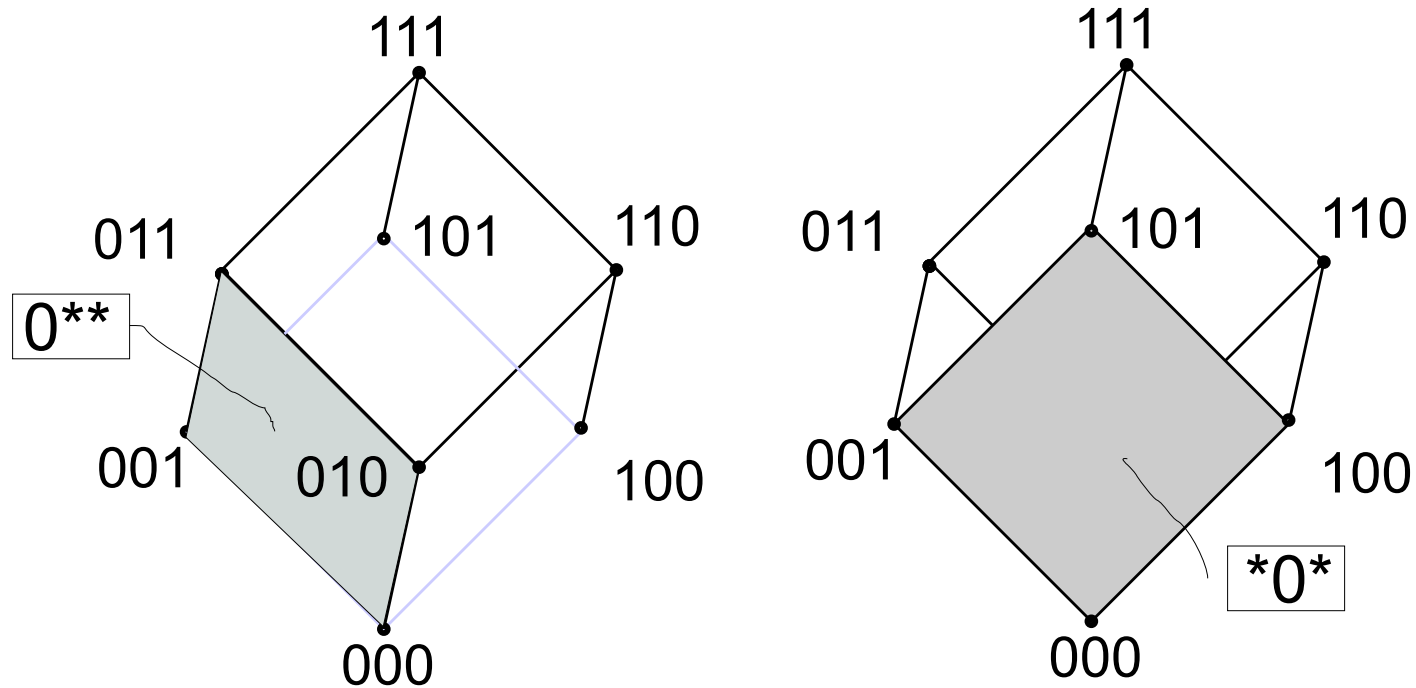
Геометрично всі такі шаблони описуються прямі – ребра куба.



Геометрична інтерпретація 5

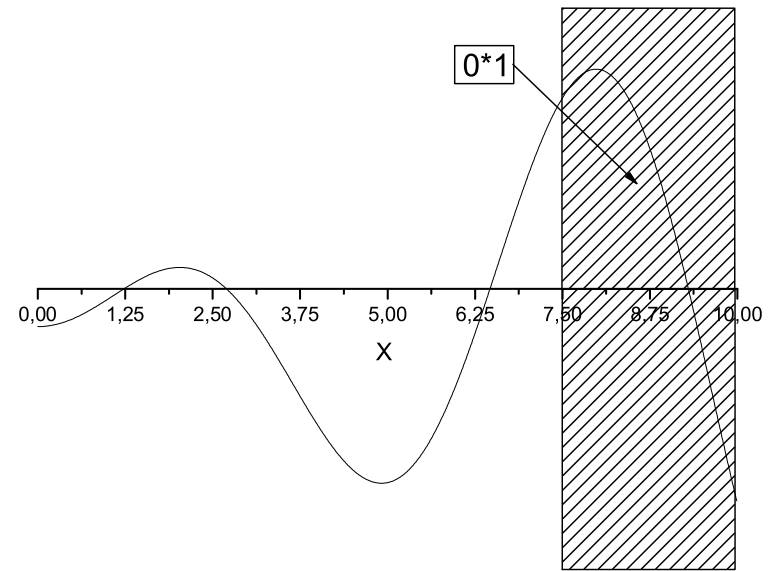
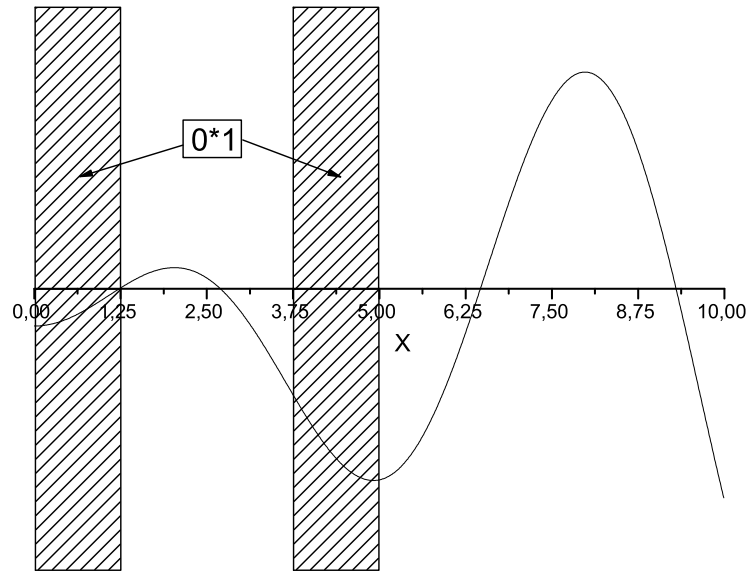
Шими 1-го порядку:

Геометрично вони описують гіперповерхні порядку, на 2 більшого від точки
– площини, і в даному випадку є гранями куба.



Шимі порядку 0 відповідає весь куб.

Інтерпретація шим в просторі параметрів



Характеристики шим – 2

Пристосованість F_S шими S :

$$F_S^t = \frac{1}{\phi_S^t} \sum_{i=1}^{\phi} F_i ; \quad l_i \in S$$

де ϕ_S^t – кількість особин, які відповідають шимі S в момент часу t .

Шима S може бути вибрана з ймовірністю:

$$P_S = \frac{F_S}{F_t} ; \quad F_t = \sum_{i=1}^N F_i$$

де F_t – пристосованість популяції.

Фундаментальна теорема ГА

$$\phi_S^{t+1} \geq \phi_S^t \frac{F_S}{\bar{F}^t} \left[1 - p_c \frac{\delta(S)}{L-1} - p_m o(S) \right]$$

де ϕ_S^{t+1} і ϕ_S^t – кількість особин, що належать шимі S відповідно в мومت часу $t + 1$ та t , $\bar{F}^t = \frac{F^t}{n}$ – середня пристосованість популяції, L – довжина генотипу.

Наслідок: Кількість особин, що належать шимам з малою характерною довжиною, малим порядком і пристосованістю, вищою від середньої по популяції, росте експоненційно, і вони швидко витісняють всіх інших. Власне це є причиною передчасної збіжності, яка безпосередньо пов'язана з FPR.

Гіпотеза будівельних блоків

Живучими є шими:

- З малим порядком – вони менш вразливі для мутацій.
- З малою характерною довжиною – для них менша ймовірність бути зруйнованими кросинговером.

Гіпотеза будівельних блоків стверджує що кросинговер дає можливість об'єднувати високоефективні малі шими в одному геномі, створюючи особини з великою пристосованістю.

З геометричної точки зору це відповідає пошуку перетинів гіперплощин з високою пристосованістю.

Приховане розпаралелення

За час, коли генетичний алгоритм опрацьовує N геномів, фактично здійснюється аналіз близько N^3 корисних шим⁶. Це – неявне розпаралелення, яке власне і служить причиною високої ефективності ГА.

⁶ Живучих і з високою пристосованістю

Параметри ГА

Набір параметрів, які необхідно правильно вибрати при практичному використанні простого ГА:

- Розмір популяції
- Ймовірність кросинговеру
- Ймовірність мутацій
- Вигляд функції пристосованості

Рівноважні ГА (Steady-State Genetic Algorithms)

На відміну від простого ГА, де нащадки повністю замінюють предків, у *рівноважному* ГА для розмноження вибирають лише частину популяції – репродуктивний інтервал (reproductive gap). Отримані особини заміщають певним чином вибраних особин з вихідної популяції.

Найпростіший випадок – за раз для розмноження вибирається одна пара.

Для таких алгоритмів стає важливим метод заміни старих членів популяції новими.

Найпростіші способи заміни:

- Замінюються найменш пристосовані члени популяції – за такої схеми швидкість збіжності значно зростає, але і зростає ймовірність виродження популяції в локальному мінімумі.
- Заміна, з ймовірністю, обернено пропорційною до пристосованості особин (inverse ranking) – дає шанс вижити погано пристосованим особинам, які все-таки можуть володіти корисною інформацією про простір пошуку.

Короткий огляд складових частин ГА – Масштабування

Масштабування цільової функції – спосіб завадити передчасній збіжності та стимулювати видоутворення.

- Лінійне масштабування $F' = aF + b$
- Сигма масштабування (Sigma scaling) –

$$F'(i, t) = \begin{cases} 1 + \frac{F(i) - \bar{F}(t)}{2\sigma(t)} & \sigma(t) \neq 0 \\ 1.0 & \sigma(t) = 0 \end{cases}$$

де $\sigma(t)$ – стандартне відхилення пристосованості популяції. Відмінне від нуля значення при $\sigma(t) = 0$ завжди залишає шанс для мало пристосованих особин залишити потомство. Застосовується для зменшення тиску відбору.

Також в конкретних задачах використовуються інші масштабування
(ступеневе $F' = F^k$; $k \approx 1.005$, тощо)

Селекція

- Випадковий вибір
- Турнірна селекція: з популяції вибираються випадковим чином 2 елементи. Генерується випадкове число $r \in [0..1]$. Якщо $r < k$ де k – певна константа (часто близька до 0.75), то для розмноження береться більш пристосована особина, інакше – менш пристосована.
- Больцманівська селекція:

$$P_{sel}(i, t) = \frac{e^{F(i)T}}{\sum_j e^{F(j)T}}$$

Температуру T зменшують в процесі еволюції для досягнення вищої точності в кінці, не допускаючи колапсу в локальний мінімум.

- Рангова селекція (rank selection): Членам популяції присвоюється ранг – значення від 1 до N згідно їхньої пристосованості, причому найкращий член має ранг N .

$$P_{sel}(i, t) = \frac{\text{rank}(i, t)}{N(N + 1)/2}$$

При великій початковій дисперсії пристосованості не дає трохи краще пристосованим заповнити своїми нащадками популяцію, до того як вона зможе оцінити простір пошуку, але забезпечує великий тиск добору при малій дисперсії – при виборі відіграє роль лише різниця в рангах а не абсолютна різниця пристосованостей.

- Групова селекція (crowding selection): пара для особини l_i вибирається так: з популяції вибираються випадковим чином F елементів (на практиці $= 2..5$). З них для розмноження вибираються найбільш подібний до l_i (найближчі в просторі пошуку).
- Елітизм (elitism) – не залежно від метода селекції і заміщення, задана кількість найбільш пристосованих членів популяції не заміщається.

Задача пошуку багатьох оптимумів. Екологічні ніші і видоутворення.

Більшість ГА з часом приводять до того, що популяція вироджується в
якомусь з мінімумів (наприклад в локальному, або одному з рівноцінних по
пристосованості глобальних)

*Способи добитися виникнення різних **видів** – субпопуляцій, які займають
різні мінімуми – **екологічні ніші**:*

- Преселекція – заміна нащадками своїх батьків, за умови що вони краще пристосовані.
- Групова поведінка (crowding) – нащадок заміняє найбільш подібного члена популяції з вибраної випадковим чином групи.

- Розподілення (sharing) – чим більше присутньо особин, подібних на дану, тим менший її шанс прийняти участь в розмноженні. Наприклад, візьмемо триангуляційну функцію розподілення:

$$Sh(d_{ij}) = 1 - \frac{d_{ij}}{\sigma_{share}}$$

Де d_{ij} – віддаль між особинами l_i, l_j в просторі пошуку, σ_{share} – певна константа, що визначає середній розмір ніш. (Якщо є вузькі, близько розташовані мінімуми, то при $\sigma_{share} > d_{min}$ вони будуть розглядатися як одна екологічна ніша!).

Тоді, трансформована пристосованість буде:

$$F'_i = \frac{F_i}{\sum_{j=1}^N Sh(d_{ij})}$$

- Модифікації розподілення – використання фенотипічної віддалі, кластерний підхід, тощо.
- Детерміністична групова поведінка (Deterministic crowding) – спарюються незалежно від їхньої пристосованості, а відбір діє на етапі заміщення, використовуючи преселекцію.
- Виключення з розгляду вже знайдених мінімумів.
- Використання субпопуляцій, кожна з яких еволюціонує незалежно але іноді обмінюється членами з іншими субпопуляціями.
- Використання турнірного заміщення.
- І. т. д.

Multi-Niche Crowding (MNC)

Всі перелічені вище алгоритми мають один суттєвий недолік – вони нездатні підтримувати різноманітність тривалий час – в процесі еволюції всі ніші крім одної "вимирають".

Алгоритм, що підтримує стабільні види необмежений час – ***MNC***.

Модифікації стандартного ГА:

- Відбір для розмноження – групова селекція.
- Заміщення – найгірші з найбільш подібних.

Цілі модифікацій:

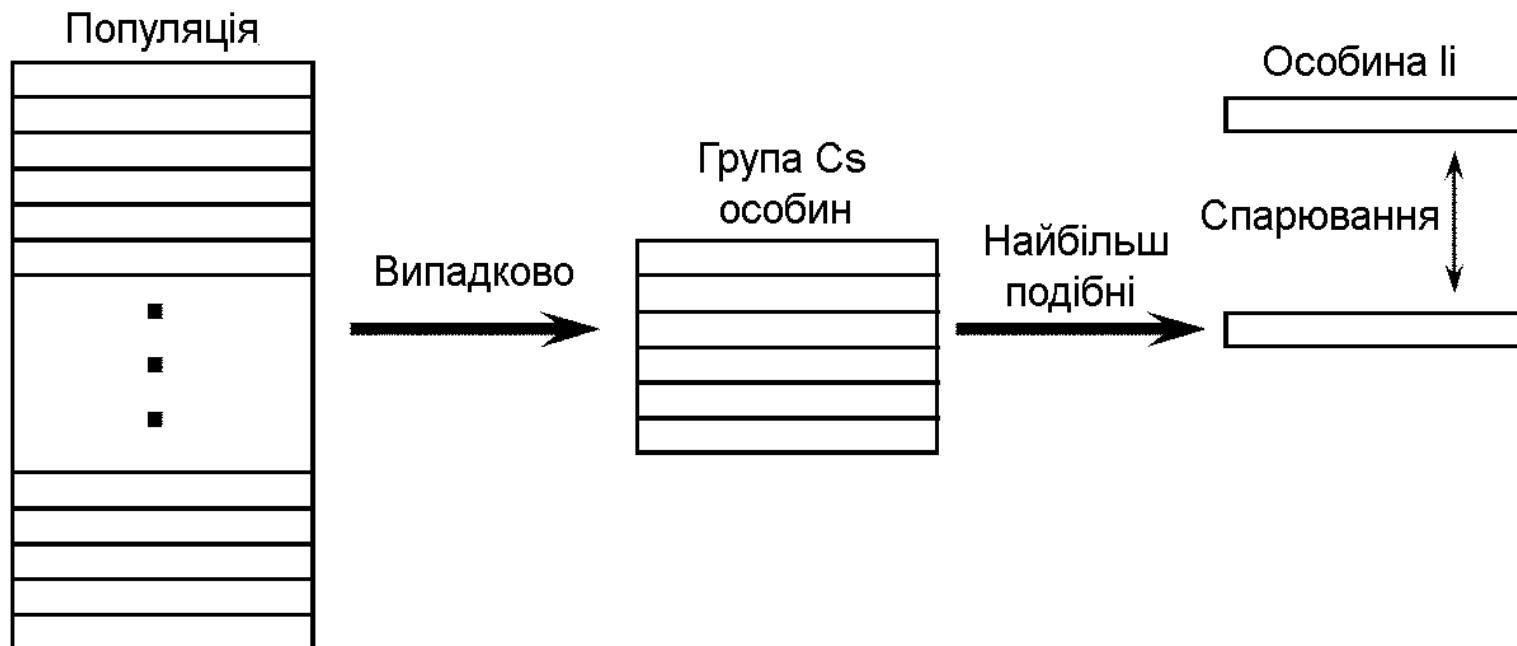
- Стимуляція спарювання в межах однієї ніші.
- Нащадки заміщають членів своєї ніші.

Основні властивості:

- Змагання відбувається як в межах ніш так і між ними.
- MNC підтримує стабільні популяції в багатьох мінімумах, в тому числі і локальних.
- *Недолік* – при малій популяції, субпопуляціям властива тенденція застрягати не в мінімумі, а десь біля нього.

Групова селекція

- Беремо початкову особину l_i з популяції (послідовно або випадково).
- Випадковим чином вибираємо C_s членів популяції.
- Для спарювання беремо особину з множини $\{C_s\}$, найближчу до l_i .



Заміщення найгірших з найбільш подібних

- Створити C_f груп з s членами в кожній.
- В кожній групі знати особину, найбільш подібну до нащадка, якого повертаємо в популяцію. З цих особин створити так-звану колективну фактор-групу.
- Замінити найменш пристосованого з фактор-групи.

