

Đào Trọng Hiếu

# GENETIC ALGORITHM

Flappy Bird

# Giải thuật di truyền

## Giới thiệu

Giải thuật di chuyền là thuật toán tìm kiếm được dựa trên sự tương tự với quá trình chọn lọc tự nhiên và thuyết tiến hóa của Charles Darwin. Giải thuật di truyền cho lời giải tốt trong nhiều bài toán tối ưu và được sử dụng rộng rãi trong nhiều

## Ứng dụng khác nhau

# Nguyên lý chung

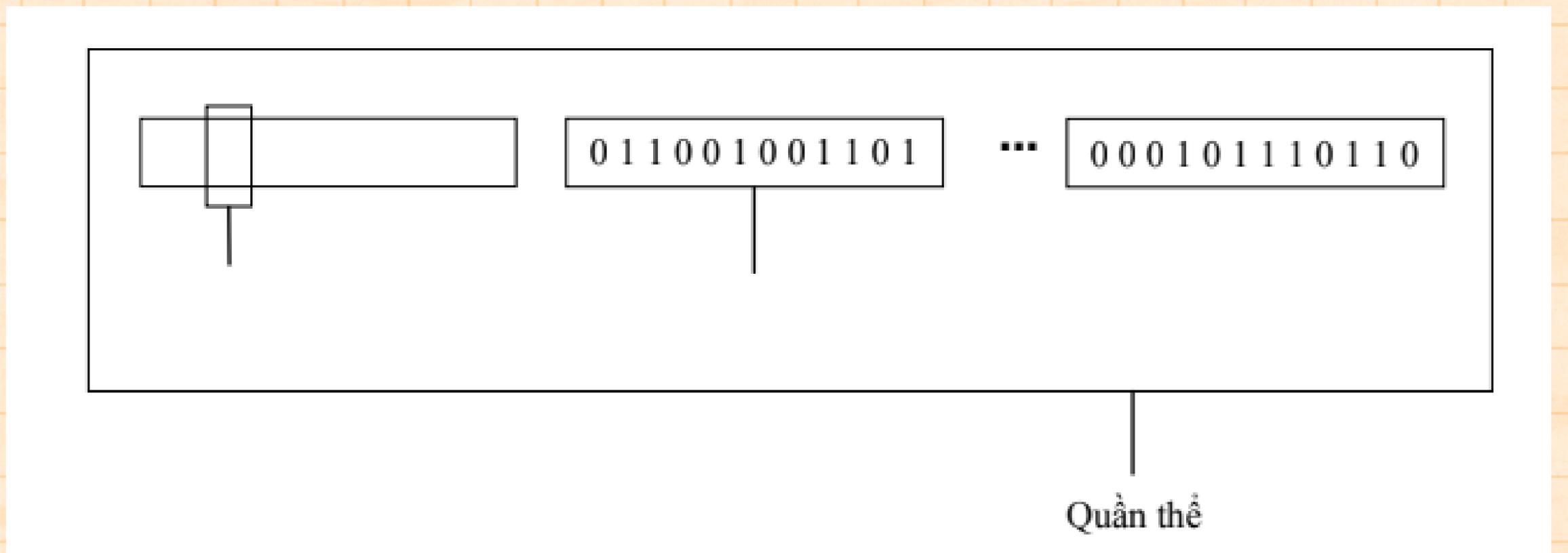
- Lời giải hay cá thể được biểu diễn như một nhiễm sắc thể
- Độ tốt của lời giải được đánh giá bởi hàm thích nghi (fitness)
- Fitness là mục tiêu của hàm tối ưu
- Mỗi thế hệ gồm những cá thể (những lời giải nhất định).

# Quy tắc

1. Các cá thể cạnh tranh với nhau. Cá thể có độ thích nghi cao hơn sẽ tạo ra nhiều hậu duệ hơn cá thể kém thích nghi

2. Gen từ các cá thể thích nghi tốt được kết hợp với nhau, nhờ đó có thể tạo ra hậu duệ tốt hơn tổ tiên
3. Từng cá thể có thể thay đổi kiểu gen của mình (Đột biến)
4. Nhờ vậy, thế hệ tiếp theo sẽ chứa cá thể thích nghi hơn với môi trường

# Ví dụ minh họa về gen, nhiễm sắc thể và quần thể



Biểu diễn lời giải và  
không gian tìm kiếm

Nhiễm sắc thể là một vec tơ có độ dài xác định, mỗi phần tử của vec tò là một gen.

## Các thao tác tạo ra quần thể mới

- Chọn lọc
- Lai ghép
- Đột biến

## Giải thuật

- Chọn lọc (selection). Chọn lọc cá thể để tham gia việc tạo ra thế hệ tiếp theo.
- Lai ghép (crossover), còn gọi là trao đổi chéo, hay tái tổ hợp.
- Đột biến (mutation). Thay đổi giá trị các gen của cá thể con vừa tạo ra với một xác suất nhất định.

Quần thể thứ  $i$

20

24415124

24

24748552

23

32752411

11

32543213

Chọn lọc

24748552

Lai ghép

24748524

Đột biến

24738524

24415124

24415152

24415152

24748552

24748511

24748511

32752411

32752452

32752456

Quần thể thứ  $i + 1$

$GA(X, f, N, c, m)$

Đầu vào: bài toán tối ưu với không gian trạng thái  $X$

hàm thích nghi  $f$

kích thước quần thể  $N$

xác suất lai ghép  $c$

xác suất đột biến  $m$

Đầu ra: lời giải (cá thể) với độ thích nghi cao

Khởi tạo: Khởi tạo ngẫu nhiên quần thể  $G$  gồm  $N$  lời giải

---

While (chưa thỏa mãn điều kiện dừng) do

1. For  $i = 1$  to  $N$  do

Tính giá trị hàm thích nghi  $f(i)$  cho cá thể thứ  $i$

2. For  $i = 1$  to  $[N/2]$  do

i. Chọn lọc: chọn 2 cá thể bố mẹ  $x$  và  $y$  từ  $G$  tuỳ theo giá trị thích nghi

ii. Lai ghép: với xác suất  $c$ , đổi chỗ đoạn gen trên  $x$  và  $y$

iii. Đột biến: với xác suất  $m$ , thay đổi giá trị các gen trên cá thể mới tạo ra

iv. Thêm cá thể mới vào quần thể mới  $G'$

3. Gán  $G \leftarrow G'$

Return: Lời giải thuộc  $G$  với giá trị thích nghi tốt nhất

# Các phương pháp thực hiện chọn lọc

Bánh xe roulette: là phương pháp thường dùng nhất. Bánh xe roulette là phương pháp chọn ngẫu nhiên hay dùng trong một trò chơi ở sòng bạc.

Theo phương pháp bánh xe roulette, xác suất chọn lời giải thứ i được tính theo công thức:

$$p(i) = \frac{f(i)}{\sum_{i=1..N} f(i)}$$

Phương pháp bánh xe roulette có thể cài đặt bằng thuật toán sau:

- + Tính tổng giá trị thích nghi của tất cả cá thể trong quần thể, gọi tổng này là S.
- + Sinh ra số ngẫu nhiên trong khoảng  $(0, S)$ , gọi số này là r.
- + Lần lượt cộng giá trị thích nghi của từng cá thể trong quần thể, gọi tổng này là s. Khi nào  $s > r$  thì dừng và trả về cá thể hiện thời.

# Thi đấu

Chọn ngẫu nhiên một đôi cá thể sử dụng phân bố xác suất đều, sau đó chọn cá thể tốt hơn trong hai cá thể đó như một cá thể cha mẹ. Thực hiện tương tự để chọn cá thể cha mẹ còn lại.

# Lựa chọn cá thể tinh hoa

Các phương pháp trên đều không đảm bảo cá thể tốt nhất được lựa chọn. Trong phương pháp lựa chọn tinh hoa, một số lượng nhất định các cá thể tốt nhất được lựa chọn trước, sau đó phần còn lại được lựa chọn theo các phương pháp ru lết hay thi đấu như ở trên. Như vậy, các cá thể tốt nhất luôn được duy trì đoạn gen của mình sang thế hệ sau và tránh làm mất lời giải tốt nhất đã tìm được.

# Giá trị xác suất

Xác suất lai ghép. Xác suất lai ghép được lựa chọn tương đối lớn, thường từ 0.5 trở lên.

Xác suất đột biến. Xác suất đột biến được lựa chọn rất nhỏ, ít khi vượt quá 0.1. Xác suất đột biến nhỏ để tránh cho thuật toán di chuyển theo kiểu ngẫu nhiên.