

THUẬT GIẢI DI TRUYỀN VÀ ỨNG DỤNG

GENETIC ALGORITHM AND ITS APPLICATION

SVTH: NGUYỄN THỊ THÚY HOÀI

Lớp: 04CCT01, Trường Đại Học Sư Phạm.

GVHD: PGS.TSKH TRẦN QUỐC CHIẾN

Khoa Tin học, Trường Đại Học Sư Phạm.

TÓM TẮT

Thuật giải di truyền (Genetic Algorithm_GA) là kỹ thuật chung giúp giải quyết vấn đề-bài toán bằng cách mô phỏng sự tiến hóa của con người hay của sinh vật nói chung (dựa trên thuyết tiến hóa muôn loài của Darwin) trong điều kiện qui định sẵn của môi trường. GA là một thuật giải và mục tiêu của GA không nhằm đưa ra lời giải chính xác tối ưu mà là đưa ra lời giải tương đối tối ưu.

ABSTRACT

Genetic Algorithm (GA) is one of search techniques in popular. The basic concept of GA is designed to simulate processes in natural system necessary for evolution, specifically those that follow the principles first laid down by Charles Darwin of survival of the fittest.

1. Mở đầu

1.1. Lý do chọn đề tài

Trong ngành khoa học máy tính, tìm kiếm lời giải tối ưu cho các bài toán là vấn đề được các nhà khoa học máy tính đặc biệt rất quan tâm.

Mục đích chính của các thuật toán tìm kiếm lời giải là tìm ra lời giải tối ưu nhất cho bài toán trong thời gian nhỏ nhất. Các thuật toán như tìm kiếm không có thông tin / vét cạn (tìm kiếm trên danh sách, trên cây hoặc đồ thị) sử dụng phương pháp đơn giản nhất và trực quan nhất hoặc các thuật toán tìm kiếm có thông tin sử dụng heuristics để áp dụng các tri thức về cấu trúc của không gian tìm kiếm nhằm giảm thời gian cần thiết cho việc tìm kiếm được sử dụng nhiều nhưng chỉ với không gian tìm kiếm nhỏ và không hiệu quả khi tìm kiếm trong không gian tìm kiếm lớn.

Tuy nhiên, trong thực tiễn có rất nhiều bài toán tối ưu với không gian tìm kiếm rất lớn cần phải giải quyết. Vì vậy, việc đòi hỏi thuật giải chất lượng cao và sử dụng kỹ thuật trí tuệ nhân tạo đặc biệt rất cần thiết khi giải quyết các bài toán có không gian tìm kiếm lớn. Thuật giải di truyền (genetic algorithm) là một trong những kỹ thuật tìm kiếm lời giải tối ưu đã đáp ứng được yêu cầu của nhiều bài toán và ứng dụng.

Hiện nay, thuật toán di truyền cùng với logic mờ được ứng dụng rất rộng rãi trong các lĩnh vực phức tạp. Thuật toán di truyền kết hợp với logic mờ chứng tỏ được hiệu quả của nó trong các vấn đề khó có thể giải quyết bằng các phương pháp thông thường hay các phương pháp cổ điển, nhất là trong các bài toán cần có sự lượng giá, đánh giá sự tối ưu của kết quả thu được. Chính vì vậy, thuật giải di truyền đã trở thành đề tài nghiên cứu thú vị và đem đến nhiều ứng dụng trong thực tiễn.

Ngày nay, GA được ứng dụng khá nhiều trong các lĩnh vực như khoa học, kinh doanh và giải trí. Đầu tiên phải kể đến là các bài toán tối ưu bao gồm tối ưu số và tối ưu tổ hợp đã sử dụng GA để tìm lời giải như là bài toán người du lịch (Travelling Salesman Problems - TSP). Ứng dụng kế tiếp của GA là thiết kế và điều khiển robo. Hầu hết các nước có ngành CNTT phát triển đã và đang rất quan tâm đến lĩnh vực thiết kế robo nhằm giúp con người tiết kiệm sức lao động và giải phóng con người thoát khỏi các công việc nguy hiểm, đặc biệt hiện nay cuộc thi “Robocon” Châu Á_ Thái Bình Dương được các nước trong khu vực rất quan tâm. Ngoài phần

cơ, để robo có thể tiến hành các hoạt động đơn giản nhất như đi, đứng... thì robo cần phải trang bị chương trình được lập trình dựa trên các thuật toán và ngôn ngữ thích hợp. Nhờ vào lịch trình được cài đặt cùng với một trí tuệ nhân tạo..., robo có thể định hướng thực hiện các hoạt động như con người. Tuy nhiên, việc tìm kiếm lời giải tốt nhất cho các hành động của robo không phải là đơn giản. Theo các nhà khoa học máy tính, thuật giải di truyền là một trong những thuật toán tối ưu giúp robo vạch lộ trình khi di chuyển. Với lý do trên, em chọn đề tài: “Thuật giải di truyền và ứng dụng”.

1.2. Đối tượng nghiên cứu

Đối tượng nghiên cứu: thuật giải di truyền và các ứng dụng

1.3. Giải pháp công nghệ

Ngôn ngữ Java

MyEclipse

2. Nội dung

2.1. Cơ sở lý thuyết

Thuật toán di truyền gồm có bốn quy luật cơ bản là lai ghép, đột biến, sinh sản và chọn lọc tự nhiên như sau:

Quá trình lai ghép (phép lai)

Quá trình này diễn ra bằng cách ghép một hay nhiều đoạn gen từ hai nhiễm sắc thể cha-mẹ để hình thành nhiễm sắc thể mới mang đặc tính của cả cha lẫn mẹ. Phép lai này có thể mô tả như sau:

Chọn ngẫu nhiên hai hay nhiều cá thể trong quần thể. Giả sử chuỗi nhiễm sắc thể của cha và mẹ đều có chiều dài là m .

Tìm điểm lai bằng cách tạo ngẫu nhiên một con số từ 1 đến $m-1$. Như vậy, điểm lai này sẽ chia hai chuỗi nhiễm sắc thể cha-mẹ thành hai nhóm nhiễm sắc thể con là m_1 và m_2 . Hai chuỗi nhiễm sắc thể con lúc này sẽ là $m_{11}+m_{22}$ và $m_{21}+m_{12}$.

Đưa hai chuỗi nhiễm sắc thể con vào quần thể để tiếp tục tham gia quá trình tiến hóa

Quá trình đột biến (phép đột biến)

Quá trình tiến hóa được gọi là quá trình đột biến khi một hoặc một số tính trạng của con không được thừa hưởng từ hai chuỗi nhiễm sắc thể cha-mẹ. Phép đột biến xảy ra với xác suất thấp hơn rất nhiều lần so với xác suất xảy ra phép lai. Phép đột biến có thể mô tả như sau:

Chọn ngẫu nhiên một số k từ khoảng $1 \leq k \leq m$

Thay đổi giá trị của gen thứ k

Đưa nhiễm sắc thể con vào quần thể để tham gia quá trình tiến hóa tiếp theo

Quá trình sinh sản và chọn lọc (phép tái sinh và phép chọn)

Phép tái sinh: là quá trình các cá thể được sao chép dựa trên độ thích nghi của nó. Độ thích nghi là một hàm được gán các giá trị thực cho các cá thể trong quần thể của nó. Phép tái sinh có thể mô phỏng như sau:

Tính độ thích nghi của từng cá thể trong quần thể, lập bảng cộng dồn các giá trị thích nghi đó (theo thứ tự gán cho từng cá thể) ta được tổng độ thích nghi. Giả sử quần thể có n cá thể. Gọi độ thích nghi của cá thể thứ i là F_i , tổng dồn thứ i là F_i . Tổng độ thích nghi là F_m

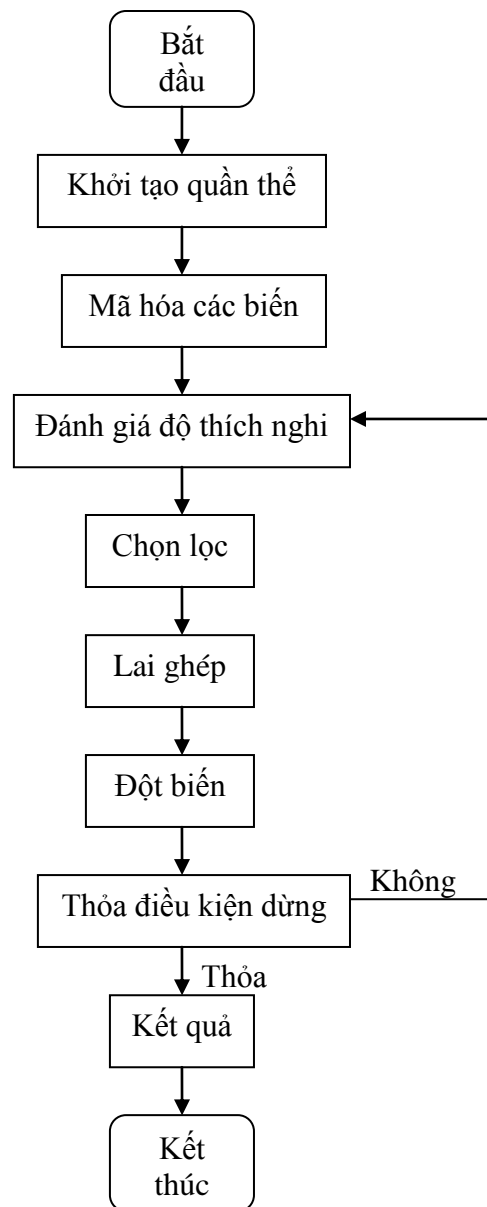
Tạo số ngẫu nhiên F có giá trị trong đoạn từ 0 đến F_m

Chọn cá thể k đầu tiên thỏa mãn $F \geq F_k$ đưa vào quần thể của thế hệ mới.

Phép chọn: là quá trình loại bỏ các cá thể xấu và để lại những cá thể tốt. Phép chọn được mô tả như sau:

Sắp xếp quần thể theo thứ tự độ thích nghi giảm dần

Loại bỏ các cá thể cuối dãy, chỉ để lại n cá thể tốt nhất.



Hình 1: Sơ đồ cấu trúc thuật toán di truyền

Cấu trúc thuật giải di truyền tổng quát

Bắt đầu

$t = 0;$

Khởi tạo $P(t)$

Tính độ thích nghi cho các cá thể thuộc $P(t);$

Khi (điều kiện dừng chưa thỏa) lặp

$t = t + 1;$

Chọn lọc P(t)

Lai P(t)

Đột biến P(t)

Hết lặp

Kết thúc

2. Các công thức của thuật giải di truyền

Tính độ thích nghi $eval(v_i)$ của mỗi nhiễm sắc thể v_i ($i=1..$ kích thước quần thể):

$$eval(v_i) = \frac{f(v_i)}{\sum_{i=1}^{kích\ thuoc\ quan\ the} f(v_i)}$$

với $f(v_i)$ là hàm mục tiêu
 Tìm tổng giá trị thích nghi của quần thể

$$F = \sum_{i=1}^{kích_thuoc_quan_the} eval(v_i)$$

Tính xác suất chọn p_i cho mỗi nhiễm sắc thể v_i

$$p_i = \frac{eval(v_i)}{\sum_{i=1}^{kích_thuoc_quan_the} eval(v_i)}$$

Tính xác suất tích lũy q_i cho mỗi nhiễm sắc thể v_i

$$q_i = \sum_{j=1}^i p_j$$

Tiến trình chọn lọc được thực hiện bằng cách quay bánh xe rulet kích thước quần thể lần. Mỗi lần chọn ra một nhiễm sắc thể từ quần thể hiện hành vào quần thể mới theo cách sau: Phát sinh một số ngẫu nhiên r trong khoảng $[0, 1]$

Nếu $r < q_1$ thì chọn nhiễm sắc thể v_1 , ngược lại chọn nhiễm sắc thể v_i ($2 \leq i \leq$ kích thước quần thể) sao cho $q_{i-1} < r \leq q_i$

3. Kết luận

3.1. Ưu điểm

Trình bày và giới thiệu những khái niệm cơ bản, cơ sở lý thuyết về thuật giải di truyền. Trên cơ sở lý thuyết, đề tài đã cài đặt các phép toán cơ bản của thuật giải di truyền nhằm phục vụ cho việc thực hiện các ứng dụng.

Sử dụng các phép toán của thuật giải di truyền để xây dựng ứng dụng cho bài toán người du lịch và bài toán vạch lộ trình đường đi cho robo

3.2. Hạn chế

Đề tài chỉ giới thiệu những kiến thức chung nhất về thuật giải di truyền, chưa đi sâu vào các vấn đề nghiên cứu tối ưu khác.

Phần ứng dụng vạch lộ trình đường đi cho robo chưa hoàn hảo. Đặc biệt là chưa giải quyết tốt việc robo tránh vật chướng ngại và kích thước quần thể thay đổi.

Giao diện chưa đẹp

3.3. Hướng phát triển

Tiếp tục nghiên cứu cơ sở lý thuyết về thuật giải di truyền

Tiếp tục phát triển ứng dụng vạch lộ trình đường đi cho robo linh hoạt hơn và tối ưu hơn.

TÀI LIỆU THAM KHẢO

[1] David A.Coley: An Introduction to Genetic Algorithm

[2] TS. Nguyễn Đình Thúc: Lập trình tiến hóa