Thông tin sinh viên:

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |
|  |  |  |

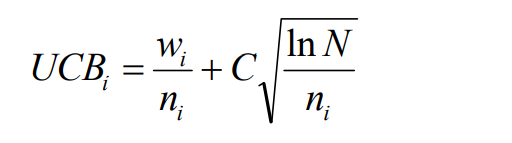
I. Giới thiệu thuật toán Monte Carlo Tree Search:

a. Giới thiệu

MCTS là phương pháp tìm kiếm dựa theo lấy mẫu dùng giả lập ngẫu nhiên để ước lượng tỷ lệ thắng thua của một trạng thái bàn cờ nhằm tìm ra nước đi tốt nhất và để cân bằng giữa việc khám phá và khai thác của tất cả các nước đi. Điểm mấu chốt của MCTS so với các phương pháp tìm kiếm truyền thống như AlphaBeta và A\* là nó không phụ thuộc vào tri thức đặc trưng của trò chơi, nói cách khác là không phụ thuộc vào hàm lượng giá trạng thái. Khi đó, MCTS có thể áp dụng vào nhiều trò chơi dạng không may rủi và thông tin trạng thái trò chơi rõ ràng sau mỗi lượt đi. Đối với các trò chơi mà khó xây đựng hàm lượng giá trạng thái tốt như cờ Vây thì việc áp dụng MCTS rất hiệu quả.

b. Cấu trúc cây tìm kiếm Monte Carlo

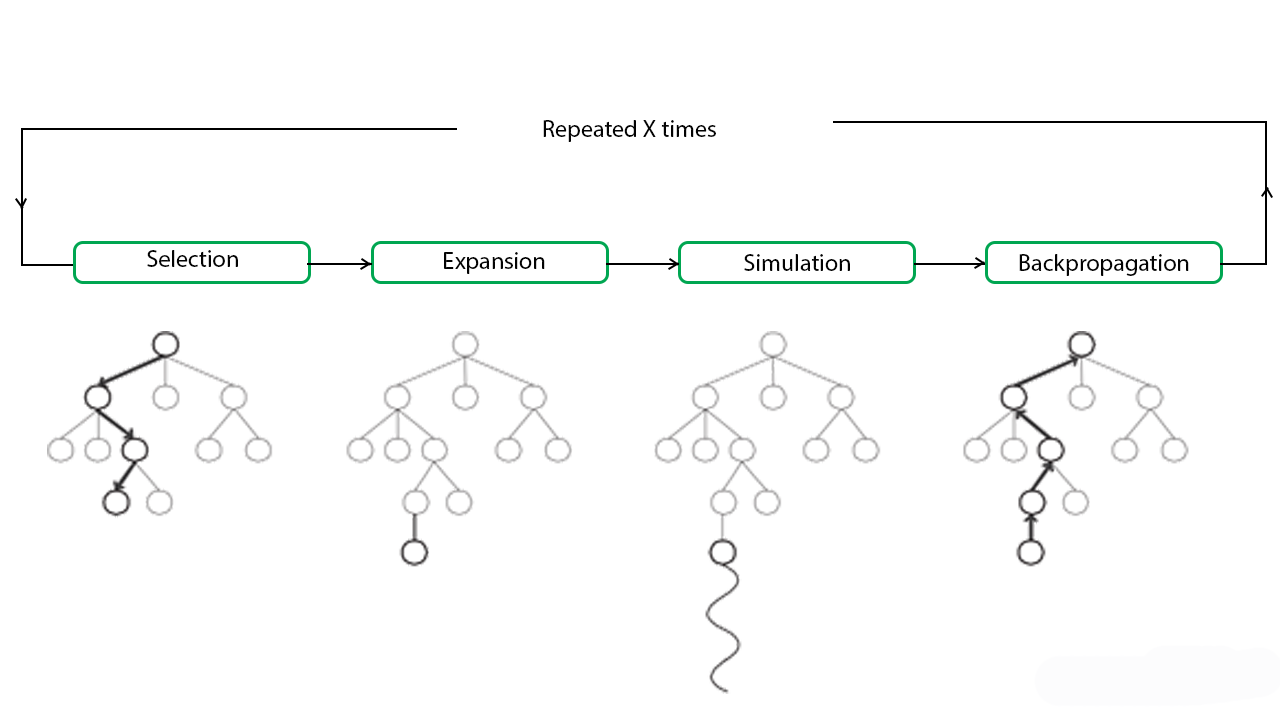
MCTS là một quá trình lặp đi lặp lại bốn bước trong một khoảng thời gian hữu hạn.

B1: Lựa chọn(Selection), thực hiện duyệt từ một nút gốc (trạng thái hiện tại) cho đến lút lá, vì vậy sẽ có nhiều hướng đi được mang ra đánh giá. Ở trong thuật toán cài đặt cho chương trình TicTacToe, chúng em sử dụng công thức UCB (cận tin cậy trên) : Với C = 1. Trong đó tỷ số wi/ni là tỷ lệ thắng của nút con i, wi là số lần thắng ván đấu giả lập có đi qua nút này, ni là số lần nút này được đi qua, N là số lần nút cha được đi qua, và C là tham số có thể điều chỉnh.

B2: Trong quá trình duyệt, khi tìm thấy nút con cũng là nút lá, MCTS sẽ chuyển sang bước mở rộng(Expansion). Một nút con mới được thêm vào cây và tính toán UCB cho nút con đó sau mỗi thực hiện lại B1. Việc mở rộng không thực hiện trừ khi kết thúc ván cờ tại nút lá.

B3: Mô phỏng (Simulation): Trong quá trình này, một mô phỏng được thực hiện bằng cách chọn nước đi hoặc chiến lược cho đến khi đạt được kết quả hoặc trạng thái xác định trước.

B4: Lan truyền ngược (Backpropagation): Sau khi xác định giá trị của nút mới được thêm vào, cây còn lại phải được cập nhật. Vì vậy, quá trình backpropagation được thực hiện, trong đó nó backpropagation từ nút mới đến nút gốc. Trong quá trình này, số lượng mô phỏng được lưu trữ trong mỗi nút sẽ tăng lên. Ngoài ra, nếu mô phỏng của nút mới dẫn đến chiến thắng, thì số lần thắng cũng tăng lên.



Cây trò chơi tăng trưởng sau mỗi lần lặp của MCTS, tăng trưởng rộng hơn và sâu hơn. Nước đi hứa hẹn là nút con nào có tỷ lệ thắng thua cao hơn được chọn trong giai đoạn chọn lựa, và rồi các cây con cũng tăng trưởng ngày càng rộng hơn và sâu hơn, và việc lấy mẫu ngày càng chính xác hơn. Sau khi kết thúc thời gian tìm kiếm, nút con nào được thăm nhiều nhất tại nút gốc sẽ được chọn để đi.

c. Ưu nhược điểm

- Ưu điểm của Tìm kiếm Cây Monte Carlo:

1. MCTS là một thuật toán đơn giản để thực hiện.
2. Monte Carlo Tree Search là một thuật toán heuristic. MCTS có thể hoạt động hiệu quả mà không cần bất kỳ kiến ​​thức nào trong lĩnh vực cụ thể, ngoài các quy tắc và điều kiện kết thúc, đồng thời có thể tìm ra các bước di chuyển của riêng mình và học hỏi từ chúng bằng cách chơi các trò chơi ngẫu nhiên.
3. MCTS có thể được lưu ở bất kỳ trạng thái trung gian nào và trạng thái đó có thể được sử dụng trong các trường hợp sử dụng trong tương lai bất cứ khi nào được yêu cầu.
4. MCTS hỗ trợ mở rộng không đối xứng của cây tìm kiếm dựa trên các trường hợp mà nó đang hoạt động.

- Nhược điểm của Tìm kiếm trên cây Monte Carlo:

1. Khi cây phát triển nhanh chóng sau một vài lần lặp lại, nó đòi hỏi một lượng lớn bộ nhớ.
2. Có một chút vấn đề về độ tin cậy với Tìm kiếm trên cây Monte Carlo. Trong một số trường hợp nhất định, có thể có một nhánh hoặc một con đường, có thể dẫn đến thua cuộc trước phe đối lập khi được triển khai cho các trò chơi theo lượt đó. Điều này chủ yếu là do số lượng lớn các tổ hợp và mỗi nút có thể không được truy cập đủ số lần để hiểu kết quả hoặc thời gian để tìm kiếm kết quả của trạng thái đó quá dài.
3. Thuật toán MCTS cần một số lượng lớn các lần lặp để có thể quyết định một cách hiệu quả đường đi hiệu quả nhất. Vì vậy, tốc độ tìm kiếm kiếm cũng là một vấn đề.

d. Time and Space Complexity

Độ phức tạp thời gian của MCTS là O(mkI/C)

Độ phức tạp không gian của MCTS là O(mk) vì trong mỗi lần lặp lại,thuật toán phải ánh xạ các trạng thái qua m\*k cụm.

Trong đó:

m là số nút con

k là số mô phỏng của một nút con

e. Ứng dụng thuật toán vào trò chơi TicTacToe

- Selection

Bắt đầu với trạng thái hiện tại, trạng thái ví dụ dưới đây. Trong giai đoạn này, thuật toán duyệt và chọn các nút con dựa vào công thức UCB đã nêu ở trên.

Diagram, schematic

Description automatically generated

- Expantion

Vì không đủ thông tin để chọn một nút nữa, nên thuật toán sẽ phát triển cây bằng cách mở các nút.

Diagram, schematic

Description automatically generated

Nếu trò chơi vẫn chưa kết thúc sẽ chuyển sang giai đoạn mô phỏng.

- Simulate

Mô phỏng trò chơi từ nút hiện tại cho đến cuối và xem ai đã thắng! Các hành động dọc theo mô phỏng được chọn ngẫu nhiên. Điều quan trọng là không thêm các trạng thái đã truy cập trong giai đoạn mô phỏng vào cây trò chơi - mục đích của mô phỏng chỉ là giúp đánh giá vị trí của nút hiện tại.

Diagram, schematic

Description automatically generated

- Backpropagation

Trong giai đoạn này, thuật toán chỉ cần cập nhật các nút đã chọn với kết quả của mô phỏng đồng thời tăng số lượt truy cập n trong mỗi nút và cập nhật số lần giành chiến thắng w chỉ cho các nút màu xám nhạt.

Diagram, schematic

Description automatically generated

Sau hàng trăm hoặc hàng nghìn lần mô phỏng, chúng ta bắt đầu có được bức tranh rõ ràng hơn về giá trị tại mỗi nút là bao nhiêu.

II. Link Demo ứng dụng trò chơi TicTacToe

II. Tài liệu tham khảo

<https://stanford.edu/~rezab/classes/cme323/S15/projects/montecarlo_search_tree_report.pdf>

<https://www.geeksforgeeks.org/ml-monte-carlo-tree-search-mcts/>

https://sj.ctu.edu.vn/ql/docgia/download/baibao-31323/12-NGUYEN%20QUOC%20HUY(17-24).pdf

http://matthewdeakos.me/2018/03/10/monte-carlo-tree-search/