TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN, ĐHQG-HCM

KHOA CÔNG NGHỆ THÔNG TIN



**BÁO CÁO ĐỒ ÁN 1**

**CÁC THUẬT TOÁN TÌM KIẾM**

Môn học: Cơ sở trí tuệ nhân tạo

Lớp: 20TN

**Danh sác sinh viên thực hiện:**

**20120081 – Nguyễn Mậu Trọng Hiếu**

**20120351 – Phạm Hữu Phúc**

**20120599 - Phù Thị Kim Trang**

**TP.HCM, tháng 10 năm 2022**

# MỤC LỤC

[MỤC LỤC 2](#_Toc117423338)

[I. Thông tin chung và phân chia công việc 4](#_Toc117423339)

[II. Mức độ hoàn thành 4](#_Toc117423340)

[1. Đánh giá tổng thể 4](#_Toc117423341)

[2. Đánh giá chi tiết từng yêu cầu 4](#_Toc117423342)

[III. Giới thiệu tổng quan 5](#_Toc117423343)

[1. Tổng quan về đồ án 5](#_Toc117423344)

[2. Tổng quan về kỹ thuật 6](#_Toc117423345)

[IV. Trình bày thuật toán 6](#_Toc117423346)

[1. Biểu diễn trong không gian trạng thái 6](#_Toc117423347)

[a. Không gian trạng thái 6](#_Toc117423348)

[b. Xây dựng đồ thị 6](#_Toc117423349)

[2. Thuật toán tìm kiếm DFS (Depth First Search) 6](#_Toc117423350)

[3. Thuật toán tìm kiếm BFS (Breadth First Search) 8](#_Toc117423351)

[4. Thuật toán tìm kiếm UCS (Uniform-Best\_cost Search) 9](#_Toc117423352)

[5. Thuật toán tìm kiếm tham lam (Greedy Best First Search) 10](#_Toc117423353)

[6. Thuật toán tìm kiếm A\* 12](#_Toc117423354)

[7. Thuật toán cho bản đồ có điểm thưởng – algo1 13](#_Toc117423355)

[8. Thuật toán cho bản đồ có điểm đón – algo2 14](#_Toc117423356)

[9. Thuật toán cho bản đồ có điểm dịch chuyển – algo3 16](#_Toc117423357)

[10. Các hàm đánh giá Heuristic cơ bản 17](#_Toc117423358)

[V. Kết quả thử nghiệm và nhận xét 18](#_Toc117423359)

[1. Hướng dẫn sử dụng chương trình 18](#_Toc117423360)

[2. Thử nghiệm 20](#_Toc117423361)

[a) Bản đồ không có điểm thưởng: 20](#_Toc117423362)

[b) Bản đồ có điểm thưởng 29](#_Toc117423363)

[c) Bản đồ có điểm đón 32](#_Toc117423364)

[VI. Cách tổ chức chương trình và ghi chú lập trình 34](#_Toc117423365)

[*a.* *Thư mục bài nộp bao gồm:* 34](#_Toc117423366)

[*b.* *Mô tả file run.sh* 35](#_Toc117423367)

[VII. Tài liệu tham khảo 35](#_Toc117423368)

# Thông tin chung và phân chia công việc

* **Bộ môn**: Cơ sở trí tuệ nhân tạo
* **Đồ án 1**: Các thuật toán tìm kiếm
* **Giảng viên lý thuyết**: Thầy Lê Hoài Bắc
* **Trợ giảng phụ trách đồ án**:
  + Nguyễn Duy Khánh ([duykhanhnguyen360@gmail.com](mailto:duykhanhnguyen360@gmail.com))
  + Nguyễn Ngọc Băng Tâm ([bangtamnguyenn@gmail.com](mailto:bangtamnguyenn@gmail.com))

|  |  |  |  |
| --- | --- | --- | --- |
| STT | MSSV | Họ tên | Nội dung công việc |
| 1 | 20120081 | Nguyễn Mậu Trọng Hiếu | Cài đặt, Hỗ trợ báo cáo |
| 2 | 20120351 | Phạm Hữu Phúc | Cài đặt, Hỗ trợ báo cáo |
| 3 | 20120599 | Phù Thị Kim Trang | Hỗ trợ cài đặt, Báo cáo |

# Mức độ hoàn thành

## Đánh giá tổng thể

Nhóm đã hoàn thành 100% yêu cầu của đồ án

## Đánh giá chi tiết từng yêu cầu

|  |  |  |
| --- | --- | --- |
| STT | Tên công việc | Độ hoàn thành |
| 1 | Mức 1a: Thiết kế đủ 5 bản đồ và hoàn thành 3 thuật toán tìm kiếm không có thông tin (bao gồm cài đặt thuật toán, báo cáo, hình vẽ minh họa). | 100% |
| 2 | Mức 1b: Hoàn thành tiếp 2 thuật toán tìm kiếm có thông tin | 100% |
| 3 | Mức 2a: Đề xuất thuật toán và các hàm heuristic phù hợp để giải quyết, thiết kế đủ 3 bản đồ và mô phỏng cách thuật toán chạy trên các bản đồ. | 100% |
| 4 | Mức 2b: Cài đặt thành công thuật toán (đảm bảo kết quả giống với mô phỏng ở mức 2a). | 100% |
| 5 | Mức 3: Đề xuất thuật toán và cài đặt thành công trường hợp có các điểm đón với tối thiểu 3 bản đồ. | 100% |
| 6 | Điểm cộng: Kịch bản nâng cấp | 100% |
| 7 | Điểm cộng: Video minh họa | 100% |
|  | Tổng điểm | 11.5/10 |

# Giới thiệu tổng quan

## Tổng quan về đồ án

Bài toán tìm đường đi ngắn nhất là một trong những bài toán phổ biến và có nhiều ứng dụng thực tế (như Google Maps, bài toán Networking, …); cũng như thường xuyên xuất hiện trong các kì thi lập trinh. Đây là một dạng bài kinh điển, được hiểu đơn giản là tìm đường đi giữa hai trạng thái hoặc nhiều trong đồ thị sao cho tổng trọng số của các cạnh tạo nên đường đi đó là nhỏ nhất.

Trong phạm vi đồ án, nhóm cần nghiên cứu và cài đặt các thuật toán tìm đường đi ngắn nhất trên một bản đồ là ma trận có kích thước cho trước với dữ liệu được đọc từ file. Các thuật toán bao gồm:

• Thuật toán tìm kiếm không có thông tin:

– Thuật toán tìm kiếm DFS (Depth First Search).

– Thuật toán tìm kiếm BFS (Breadth First Search).

– Thuật toán tìm kiếm UCS (Uniform-Best\_cost Search).

• Thuật toán tìm kiếm có thông tin:

– Thuật toán tìm kiếm tham lam (Greedy Best First Search).

– Thuật toán tìm kiếm A\*.

* **Mục tiêu:** Cài đặt các thuật toán tìm kiếm đường đi từ vị trí xuất phát đến vị trí đích (vị trí thoát khỏi mê cung) cho tác nhân, và nếu tác nhân di chuyển vào ô điểm thưởng sẽ được giảm chi phí đường đi.
* **Yêu cầu ràng buộc:**

+ Tác nhân chỉ có thể di chuyển lên, xuống, trái, phải với chi phí bằng nhau.

+ Vị trí thoát khỏi mê cung là vị trí duy nhất trên biên của các bức tường mà tác nhân có thể thoát khỏi mê cung.

+ Đối với các ô trong bản đồ có thể di chuyển đến, chỉ có thể đi qua tối đa 1 lần.

+ Trường hợp bản đồ có điểm đón thì tác nhân bắt buộc phải đi qua tất cả các điểm đón.

+ Đối với bản đồ chứa ô dịch chuyển, khi tác nhân chạm vào ô dịch chuyển (teleport), tác nhân buộc phải bị dịch chuyển.

## Tổng quan về kỹ thuật

* Ngôn ngữ lập trình: Python
* Phần mềm lập trình: Visual Studio Code
* Sử dụng thư viện Pygame để thực hiện các công việc đồ họa
* Yêu cầu kỹ thuật: Python 3.5 trở lên, bao gồm công cụ cài đặt pip
* Có file requirements.txt để thuận tiện cài đặt thư viện thông qua pip

# Trình bày thuật toán

## Biểu diễn trong không gian trạng thái

### Không gian trạng thái

* Các trạng thái: Ứng với mỗi vị trí (x, y) trên ma trận
* Hành động: Di chuyển lên, xuống, trái, phải
* Mô hình chuyển tiếp: Cập nhật vị trí
* Mục tiêu: Di chuyển từ vị trí xuất phát đến vị trí đích

Kích thước không gian trạng thái của bài toán: M \* N.

### Xây dựng đồ thị

Với ma trận có kích thước M \* N, mỗi ô trong ma trận tương ứng với mỗi trạng thái của đồ thị. Ta sử dụng mảng 2 chiều để lưu giá trị của tất cả trạng thái có trong đồ thị.

Đối với mỗi trạng thái sẽ có tối đa 4 ô có thể di chuyển đến bằng hành động di chuyển lên, xuống, trái, phải. Biết rằng chi phí di chuyển đến các ô này là như nhau, đều bằng 1.

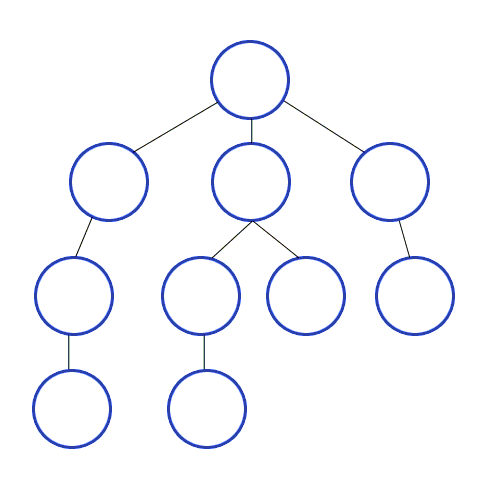
## Thuật toán tìm kiếm DFS (Depth First Search)

***Tìm kiếm theo chiều sâu*** là phương pháp tìm kiếm mà ưu tiên tìm kiếm những trạng thái xa với trạng thái xuất phát. Thuật toán DFS mở rộng tìm kiếm theo chiều sâu, nghĩa là thuật toán tìm đường theo hướng phát triển một nhánh duy nhất. Xét trên đồ thị ít vật cản hoặc không có vật cản, thuật toán sẽ tìm được đường đi hợp lệ với số lần thử ít hơn các thuật toán còn lại.

***Ý tưởng thuật toán*:** Từ trạng thái xuất phát ban đầu ta sẽ tìm được các trạng thái kề và duyệt qua các trạng thái này. Sau đó ta tìm các trạng thái kề của trạng thái vừa xét và tiếp tục duyệt đến khi đi qua hết tất cả các trạng thái có thể duyệt. Lưu ý, ta đi qua sâu vào từng nhánh của mỗi trạng thái đang xét. Thuật toán dừng khi đi qua hết tất cả các trạng thái có thể đi qua.

***Thuật toán tìm kiếm sâu:***

* Bước 1. Cho trạng thái xuất phát vào open. Lưu ý, open hoạt động theo quy tắc LIFO
* Bước 2. Nếu open rỗng thì tìm kiếm thất bại, kết thúc việc tìm kiếm.
* Bước 3. Lấy trạng thái đầu x trong open ra. Cho x vào closed.
* Bước 4. Nếu x là trạng thái đích thì tìm kiếm thành công, kết thúc việc tìm kiếm.
* Bước 5. Tìm tất cả các trạng thái con của x không thuộc open và closed cho vào đầu của open 6. Trở lại bước 2.



Hình : Minh họa thuật toán DFS (nguồn Wikipedia)

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

***Đánh giá thuật toán:***

Tổng số trạng thái cần xét tối đa là: 1 + b + b2+…+bd

Tổng số trạng thái tối đa trong open là: bd

Trong đó:

+ b: số nhánh tối đa của 1 trạng thái

+ d: độ sâu

Như vậy độ phức tạp về thời gian của thuật toán là O(bd ) và không gian (bộ nhớ) của thuật toán là O(bd).

Ta thấy thuật toán sẽ duyệt qua tất cả các trạng thái trong trường hợp đồ thị liên thông. Tuy nhiên, đây là thuật toán tìm kiếm mù, mang tính chất vét cạn và sẽ kém hiệu quả nếu tập trạng thái quá lớn. Các thuật toán tìm kiếm không thông tin không có thêm thông tin về trạng thái hoặc không gian tìm kiếm ngoài cách đi qua cây, vì vậy nó còn được gọi là tìm kiếm mù.

## Thuật toán tìm kiếm BFS (Breadth First Search)

***Tìm kiếm theo chiều rộng*** là phương pháp tìm kiếm mà ưu tiên những trạng thái gần với trạng thái xuất phát.

***Thuật toán tìm kiếm rộng:***

* Bước 1. Cho trạng thái xuất phát vào open.
* Bước 2. Nếu open rỗng thì tìm kiếm thất bại, kết thúc việc tìm kiếm. Ở giải thuật này, open hoạt động theo quy tắc FIFO.
* Bước 3. Lấy trạng thái đầu trong open ra và gọi đó là x.
* Bước 4. Nếu x là trạng thái đích thì tìm kiếm thành công, kết thúc việc tìm kiếm.
* Bước 5. Tìm tất cả các trạng thái con của x và cho vào cuối của open
* Bước 6. Trở lại bước 2

|  |  |
| --- | --- |
| Hình : Minh họa thuật toán BFS (nguồn Wikipedia) |  |
|  |  |

***Đánh giá thuật toán:***

Tổng số trạng thái cần xét tối đa là: 1 + b + b2+…+bd .

Trong đó:

+ b: số nhánh tối đa của 1 trạng thái

+ d: độ sâu

Như vậy độ phức tạp về thời gian và không gian (bộ nhớ) của thuật toán là O(bd)

Thuật toán sẽ duyệt qua tất cả các trạng thái nếu đồ thị liên thông. Tương tự như DFS, BFS cũng là thuật toán tìm kiếm mù, mang tính chất vét cạn và cũng gặp hạn chế khi tập trạng thái lớn.

Với những ưu, nhược điểm khác nhau của DFS và BFS thì việc áp dụng các giải thuật này để phục vụ cho việc giải quyết các yêu cầu bài toán phù hợp, trong đó:

* BFS: dùng để tìm đường đi ngắn nhất trên đồ thị.
* DFS: dùng để kiểm tra có tồn tại đường đi từ trạng nguồn đến trạng thái đích được hay không?

## Thuật toán tìm kiếm UCS (Uniform-Best\_cost Search)

***Thuật toán UCS*** là một thuật toán duyệt và tìm kiếm trên một cấu trúc cây, hoặc đồ thị có trọng số (chi phí). Việc tìm kiếm bắt đầu tại nút gốc và tiếp tục bằng cách duyệt các nút tiếp theo với trọng số hay chi phí thấp nhất tính từ nút gốc. Thuật toán sẽ tiếp tục duyệt và kết thúc khi gặp trạng thái đích.

***Thuật toán tìm kiếm chi phí cực* *tiểu*:**

Bước 1. Cho đỉnh xuất phát vào open.

Bước 2. Nếu open rỗng thì tìm kiếm thất bại, kết thúc việc tìm kiếm.

Bước 3. Lấy đỉnh đầu trong open ra và gọi đó là x. Cho x vào closed

Bước 4. Nếu x là đỉnh đích thì tìm kiếm thành công, kết thúc việc tìm kiếm.

Bước 5. Tìm tất cả các đỉnh con của x không thuộc open và closed cho vào open theo thứ tự tăng dần về khoảng cách từ đỉnh xuất phát.

Bước 6. Trở lại bước 2

Ảnh có chứa bàn

Mô tả được tạo tự động

Hình . Minh họa thuật toán tìm kiếm UCS (Nguồn Wikipedia)

***Đánh giá thuật toán:***

+ Luôn luôn tìm được đường đi từ trạng thái nguồn đến trạng thái đích, đảm bảo tìm được đường đi có “chi phí thấp nhất”.

+ Hiệu quả với bản đồ nhiều vật cản như trong mê cung, ngược lại đối với bản đồ không hoặc ít vật cản thì kém hiệu quả.

+ Khá tốn chi phí về mặt bộ nhớ và thời gian thực thi.

## Thuật toán tìm kiếm tham lam (Greedy Best First Search)

Để khắc phục tính kém hiệu quả của các thuật toán mù, ta tìm hiểu đến phương pháp tìm kiếm có thông tin sử dụng hàm heuristic để đánh giá, trước hết là thuật toán Greedy Best First Search (Greedy BFS). Từ “tham lam” trong thuật toán hiểu là chọn giải pháp tối ưu tại mỗi bước đi với hi vọng tìm được giải pháp tối ưu cho cả bài toán.

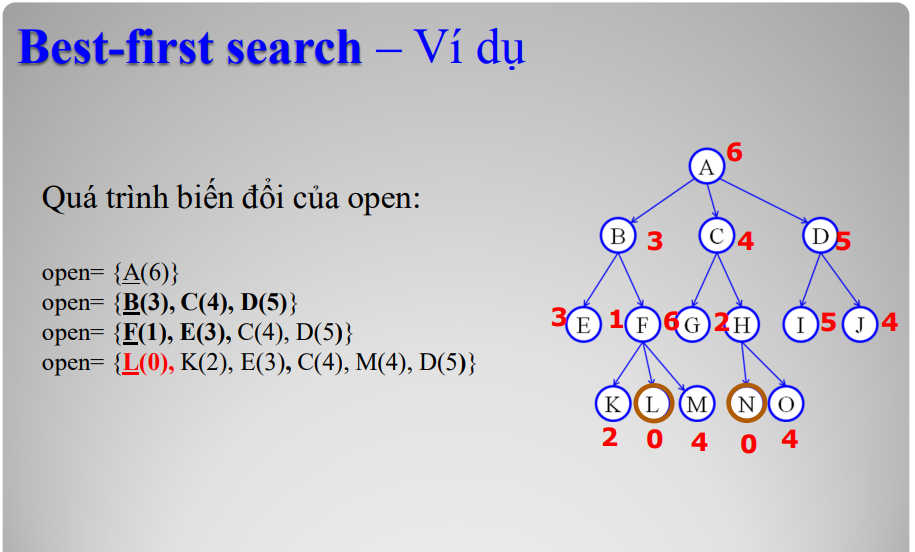
***Tư tưởng thuật toán***: Lựa chọn giải pháp nào được cho là tốt nhất ở thời điểm hiện tại và sau đó giải bài toán con nảy sinh từ việc thực hiện lựa chọn vừa rồi. Lựa chọn của thuật toán tham lam có thể phụ thuộc vào các lựa chọn trước đó. Nhưng nó không thể phụ thuộc vào một lựa chọn nào trong tương lai hay phụ thuộc vào lời giải của các bài toán con.

Thuật toán tiến triển theo kiểu thực hiện các chọn lựa theo một vòng lặp, cùng lúc đó thu nhỏ bài toán đã cho về một bài toán con nhỏ hơn. Đây cũng là sự khác biệt giữa thuật toán so với quy hoạch động.

Giải thuật tham lam quyết định sớm và thay đổi đường đi thuật toán theo quyết định đó, và không xét lại các quyết định cũ, vì vậy có thể là một thuật toán không tối ưu.

Thuật toán đánh giá chi phí các ô chỉ dựa trên chi phí từ ô đó đến ô đích:

|  |
| --- |
| f(n) = h(n) với f(n) là hàm đánh giá và h(n) là hàm heuristic. |



Hình . Mô tả thuật toán tìm kiếm tham lam

***Thuật toán tìm kiếm tham lam:***

Bước 1. Cho đỉnh xuất phát vào open.

Bước 2. Nếu open rỗng thì tìm kiếm thất bại, kết thúc việc tìm kiếm.

Bước 3. Lấy đỉnh đầu trong open ra và gọi đó là x. Cho x vào closed

Bước 4. Nếu x là đỉnh đích thì tìm kiếm thành công, kết thúc việc tìm kiếm.

Bước 5. Tìm tất cả các đỉnh con của x không thuộc open và closed cho vào open theo thứ tự tăng dần về khoảng cách ước lượng đến đích.

Bước 6. Trở lại bước 2.

***Đánh giá thuật toán:***

+ Tận dụng ưu điểm của BFS và DFS, Greedy BFS có thể chuyển đổi giữa 2 giải thuật trên, vì thế thuật toán cũng hiệu quả hơn nhưng độ hiệu quả cũng phụ thuộc lớn vào hàm Heuristic. Điểm đặc biệt là Greedy tập trung tối ưu hóa thời gian tìm đường đến đích.

+ Đường đi tìm được không đảm bảo có chi phí thấp nhất.

+ Hiệu quả với bản đồ trống, không nhiều vật cản; không hiệu quả với bản đồ nhiều vật cản và chật hẹp.

+ Tốn nhiều thời gian nếu bản đồ phức tạp nhưng tốn ít bộ nhớ

## Thuật toán tìm kiếm A\*

***Thuật toán A\*:***

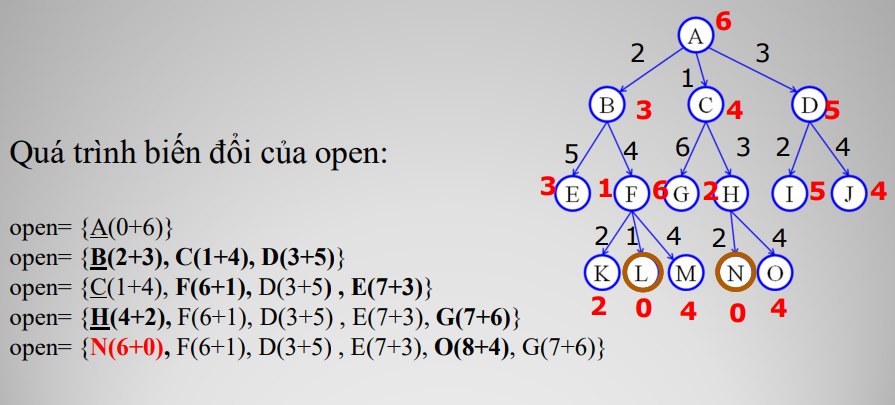
Thuật toán A\* là cũng là một trong những thuật toán Tìm kiếm có thông tin. A\* tìm đường đi từ một đỉnh hiện tại đến đỉnh đích có sử dụng hàm để ước lượng khoảng cách hay còn gọi là hàm Heuristic.

Thuật toán đánh giá chi phí các ô dựa trên tổng chi phí từ nút gốc đến nút đó và chi phí từ nút đó đến đích:

|  |
| --- |
| f(n) = g(n) + h(n) với f(n) là hàm đánh giá, g(n) là hàm chi phí thực và h(n) là hàm heuristic chi phí ước lượng. |

Biết rằng f(x) thường có giá trị càng thấp thì độ ưu tiên càng cao.

Điểm khác biệt giữa A\* và Greedy BFS là A\* tính khoảng cách đã đi qua, nên làm cho A\* tối ưu hơn và luôm tìm được đường đi ngắn nhất nếu tồn tại. Tuy nhiên, A\* không đảm bảo chạy nhanh hơn các thuật toán khác đơn giản hơn.



Hình . Mô tả thuật toán A\* (Nguồn slide thầy Phạm Minh Tuấn)

***Tư tưởng thuật toán:*** Xuất phát từ một ô và tính chi phí ước lượng của tất cả các ô lân cận (lưu ý chỉ tiếp tục duyệt đến ô kề có chi phí ước lượng tốt nhất). Lặp lại quá trình trên cho đến khi gặp ô đích hoặc không còn ô nào có thể đi.

***Thuật toán A\*:***

Bước 1. Cho đỉnh xuất phát vào open.

Bước 2. Nếu open rỗng thì tìm kiếm thất bại, kết thúc việc tìm kiếm.

Bước 3. Lấy đỉnh đầu trong open ra và gọi đó là x. Cho x vào closed

Bước 4. Nếu x là đỉnh đích thì tìm kiếm thành công, kết thúc việc tìm kiếm.

Bước 5. Tìm tất cả các đỉnh con của x không thuộc open và closed cho vào open theo thứ tự tăng dần đối với hàm f(n)=g(n)+h(n).

Bước 6. Trở lại bước 2.

***Đánh giá thuật toán:***

+ Thuật toán chắn chắn tìm được đường đi nếu có, nhưng đường đi tìm được chưa chắc là ngắn nhất (Phụ thuộc vào hàm Heuristic).

+ Hiệu quả với bản đồ trống, không nhiều vật cản và không hiệu quả với bản đồ nhiều vật cản và chật hẹp.

+ Tốn thời gian nếu bản đồ phức tạp nhưng tốn ít bộ nhớ.

+ Độ tối ưu của thuật toán phụ thuộc lớn vào hàm Heuristic.

## Thuật toán cho bản đồ có điểm thưởng – algo1

Thuật toán do chúng em đề xuất có tư tưởng giống với thuật toán A\*. Điểm khác biệt nằm trong hàm đánh giá heuristic. Với mọi đỉnh n được duyệt đến, ta không chỉ quan tâm đến f(n) = g(n) + h(n) mà còn quan tâm đến các điểm thưởng. Thực hiện duyệt qua lần lượt các điểm thưởng t, ta quan tâm đến chi phí heuristic đi từ điểm n đang xét đi đến t, cộng với giá trị thưởng tại t, rồi chi phí heuristic đi từ t đến điểm kết thúc. Tóm lại ta có công thức như sau:

|  |
| --- |
| f(n) = g(n) + min(h(n), h(n, t) + award(t) + h(t))  với mọi t là điểm thưởng chưa được duyệt, award(t) là  điểm thưởng của t , h(t) là chi phí heuristic từ t đến đích,  h(n, t) là chi phí heuristic đi từ n đến t |

Biết rằng f(x) có giá trị càng thấp thì độ ưu tiên càng cao.

***Đánh giá thuật toán:***

+ Thuật toán chắn chắn tìm được đường đi nếu có, nhưng đường đi tìm được chưa chắc là có chi phí tối ưu nhất.

+ Độ phức tạp về thời gian của thuật toán chính là độ phức tạp thời gian của thuật A\* nhân với số lượng điểm thưởng. Tuy nhiên trong giới hạn đồ án này, số điểm thưởng nằm trong khoảng vài chục và cao nhất là hàng trăm nên với giải pháp này là có thể chấp nhận được.

+ Độ tối ưu của thuật toán phụ thuộc lớn vào hàm Heuristic.

## Thuật toán cho bản đồ có điểm đón – algo2

Về hướng tiếp cận, chúng em trả lời hai câu hỏi cho bài toán bao gồm **(1)** đường ngắn nhất giữa các điểm đón với nhau là gì, **(2)** thứ tự thăm các điểm đón như thế nào là tối ưu.

***Đối với câu hỏi (1)***, nhóm tận dụng thuật toán từ mức level 1, cụ thể là sử dụng thuật toán UCS để tìm đường đi ngắn nhất từ 1 điểm đón (hoặc điểm bắt đầu) đến các điểm đón (và điểm kết thúc) của bản đồ. Về việc chọn thuật toán UCS, nhóm nhận xét đây là thuật toán uninformed search dễ cài đặt, ổn định, đảm bảo tìm đường đi luôn tối ưu và vẫn giữ được tính đúng đắn nếu có điểm thưởng trong bản đồ. Về cách lưu truy vết đường đi, nhóm sử dụng tư tưởng: xét đường đi từ *i→j*, gọi *k* là ô ngay trước *j* trong đường đi (tức là *i→ …→ k →j*), ta sẽ lưu truy vết *trace[i][j] = k* và tiến hành đệ quy ngược lại (từ *i→k* và tiếp tục như thế) đến khi đỉnh đầu trùng đỉnh cuối thì dừng. Như vậy, chúng em đã trả lời được câu hỏi thứ nhất và tiến hành tìm lời giải cho câu hỏi còn lại.

***Đối với câu hỏi (2)***, bằng cách sử dụng kết quả phía trên, nhóm xây dựng một đồ thị gồm:  
+ Tập đỉnh chứa: đỉnh bắt đầu, đỉnh kết thúc (đường ra bản đồ) và danh sách các đỉnh đón.  
+ Tập cạnh chứa: mỗi cạnh là chi phí của đường đi ngắn nhất từ một đỉnh đến đỉnh còn lại.   
Thông qua dựng đồ thị như trên, nhóm thấy rằng đây là bài toán tương tự với Travelling salesman problem (TSP), khác ở điểm: bài toán tìm đường đi với điểm đầu cố định là vị trí bắt đầu, điểm cuối cố định là vị trí ra khỏi mê cung, và giữa hai điểm này sẽ là một hoán vị của các điểm đón sao cho chi phí tổng là ít nhất. Với các cách tiếp cận như brute force hay sử dụng tìm kiếm theo chiều sâu (DFS), chương trình lại tỏ ra không hiệu quả nếu như kích thước số điểm đón lớn (≥ 10). Do đó, nhóm chọn cài đặt Genetic Algorithm để giải quyết bài toán NP này. Nhìn chung, các bước thực hiện được tóm tắt như sau:

* Bước 1. Khởi tạo ngẫu nhiên quần thể với kích thước *initial\_size* ban đầu bằng công thức:  
  Điểm bắt đầu + hoán vị của danh sách các điểm đón + điểm kết thúc.
* Bước 2: Với mỗi thế hệ thực hiện:
  + Bước 2.1. Tính độ tương thích ứng của mỗi “đối tượng”, là tổng độ dài của toàn bộ đường đi.
  + Bước 2.2. Chọn lọc tự nhiên bằng cách giữ lại các đối tượng tốt nhất, số lượng là *size\_for\_best.* Tuy nhiên, ta vẫn giữ lại 10-20% các phần tử còn lại.
  + Bước 2.3. Nhân đôi quần thể bằng cách lai tạo ra thế hệ kế tiếp của quần thể bằng cách:

Ghép một đoạn gene liên tục của “cha” với các “gene” còn lại lấy từ mẹ.  
Vd:

Cha: BĐ 0 5 3 1 2 4 KT (Đoạn gene được lấy là 3 1 2)

Mẹ: BĐ 0 4 1 5 3 2 KT (Các gene lấy từ mẹ là 0 4 5)

Con: BĐ 0 4 3 1 2 5 KT   
(Giữ đúng thứ tự các gen lấy từ cha vào con, sau đó điền các gen còn lại chưa có theo thứ tự gen của mẹ)

* + Bước 2.4. Đột biến cho quần thể với tỉ lệ là *mutation\_rate*. Việc đột biến xảy ra bằng cách hoán đổi 2 vị trí hai gene bất kỳ trong chuỗi gene hiện tại (không bao gồm điểm đầu và điểm cuối). Tương ứng với mỗi “đối tượng”, ta tạo ra phiên bản đột biến tương ứng, tuy nhiên nếu trước đó đối tượng vẫn tốt, ta giữ lại phiên bản trước đó cho quần thể.

**Về các thông số:**

+ *n\_generations*: 20000 là số lượng vòng đời.

+ *initial\_size*: 20.

+ *size\_for\_best:* 10.

+ *mutation\_rate*: 0.3.

**Giải thích:**

+ Với số điểm đón nhiều nhất là 25, chúng em thử nghiệm và nhận thấy số lượng tốt nhất cần giữ lại giới hạn là 10, và giữ lại ngẫu nhiên 10%-20% các phần tử kém, mục đích để tránh bị rơi vào vùng cực tiểu địa phương. Số lượng phần tử trong mỗi thế hệ giữ ở mức < 50 phần tử.

+ Do hạn định về thời gian nên nhóm chọn 20000 là số lượng vòng đời cần thiết.

+ Về tỉ lệ xảy ra biến dị, nhóm nhận thấy cần thay đổi hoán vị các vị trí gene nhiều để tạo ra đa dạng hoán vị, do đó cần khởi tạo *mutation\_rate* ở mức cao vừa phải. Sau khi tham khảo từ [2] và thử nghiệm, nhóm nhận thấy 0.3 phù hợp với bài toán và cho kết quả tốt.

***Đánh giá thuật toán:***

+ Thuật toán tìm được đường đi nếu có, kết quả đường đi không đảm bảo là tối ưu.

+ Độ phức tạp thời gian của thuật toán là tổng độ phức tạp của hai giai đoạn:

* Giai đoạn 1: tìm đường đi ngắn nhất cho mỗi đỉnh đón và đỉnh bắt đầu bằng thuật toán UCS. Như vậy độ phức tạp ở giai đoạn này là

(số lượng đỉnh đón + 1) \* độ phức tạp thuật toán UCS.

* Giai đoạn 2: Gọi *n* là số phần tử trung bình của mỗi thế hệ (~50), *m* là chiều dài đoạn gene (<=25). Bước thực hiện tốn chi phí nhất trong mỗi thế hệ nằm ở bước chọn lọc do phải duyệt qua mỗi đối tượng để tính chiều dài đường đi - O(n \* m) và sắp xếp lại để chọn phần tử tốt nhất – O[nlog(n)]. Như vậy, với *N\_GEN* là số lượng thế hệ, thì độ phức tạp theo Big-O là:

O(N\_GEN \* n \* (m + logn))

+ Quá trình lai tạo (bước 2.3.), đột biến và việc giữ lại các phần tử cũ tối ưu sau đột biến (bước 2.4.) đều tương ứng với quá trình tìm kiếm cực tiểu.

+ Ngoài ra, quá trình đột biến còn có tác dụng hỗ trợ việc “thoát khỏi” cực tiểu địa phương. Tuy nhiên, càng về các thế hệ sau, tác dụng này của quá trình đột biến bị hạn chế do sự xuất hiện của nhiều phần tử ở vùng cục bộ địa phương, nếu không có sự thay đổi đáng kể, đột biến chỉ đóng góp vào quá trình tìm kiếm cực tiểu.

+ Thông qua thử nghiệm, nhóm nhận thấy, kết quả của thuật toán cho ra là tối ưu với trường hợp số lượng điểm đón nhỏ (<=10) và gần tối ưu nếu số lượng điểm đón lớn (>10).

## Thuật toán cho bản đồ có điểm dịch chuyển – algo3

Chúng em quy định đây là dịch chuyển (teleport) 2 chiều. Nếu từ (i,j) đi đến (u,v) được thì từ (u,v) cũng đến được (i,j). Ngoài ra khi chạm vào điểm teleport thì bắt buộc dịch chuyển. Chi phí để teleport là 0.

Thuật toán do chúng em đề xuất có tư tưởng giống với thuật toán A\*. Điểm khác biệt nằm trong hàm đánh giá heuristic. Cũng như algo1 – dành cho bản đồ có điểm thưởng, ta không chỉ quan tâm đến f(n) = g(n) + h(n) mà còn quan tâm đến các điểm teleport. Thực hiện duyệt qua lần lượt các điểm teleport t, ta quan tâm đến chi phí heuristic đi từ điểm n đang xét đi đến t, cộng với chi phí heuristic đi từ t’ (điểm teleport từ t dịch chuyển tới) đến điểm kết thúc. Với mọi đỉnh n được duyệt đến, ta kiểm tra xem điểm đó có phải là điểm teleport không. Nếu không ta thực hiện như thuật toán A\* bình thường với hàm heuristic trình bày như trên. Tuy nhiên nếu điểm đó là điểm teleport và vì ta quy ước bắt buộc dịch chuyển, ta cần phải lưu lại truy vết điểm trước t’ là t. Sau đó thực hiện việc đánh giá và cập heuristic như bình thường nhưng mà là cho điểm t’ chứ không phải t (vì ta đã dịch chuyển rồi). Tóm lại ta có công thức heuristic như sau:

|  |
| --- |
| f(n) = g(n) + min(h(n), h(n, t) + h(t’))  với mọi t là điểm teleport chưa được duyệt, t’ là  điểm dịch chuyển từ t , h(t’) là chi phí heuristic từ t’  đến đích, h(n, t) là chi phí heuristic đi từ n đến t |

Biết rằng f(x) có giá trị càng thấp thì độ ưu tiên càng cao.

***Đánh giá thuật toán:***

+ Thuật toán chắn chắn tìm được đường đi nếu có, nhưng đường đi tìm được chưa chắc là có chi phí tối ưu nhất.

+ Độ phức tạp về thời gian của thuật toán chính là độ phức tạp thời gian của thuật A\* nhân với số lượng điểm dịch chuyển. Tuy nhiên trong giới hạn đồ án này, số điểm thưởng nằm trong khoảng vài chục và cao nhất là hàng trăm nên với giải pháp này là có thể chấp nhận được.

+ Độ tối ưu của thuật toán phụ thuộc lớn vào hàm Heuristic.

## Các hàm đánh giá Heuristic cơ bản

Trong các thuật toán được dùng ở trên đều sử dụng các hàm heuristic hoặc từ các hàm heuristic cơ bản phát triển lên các hàm đánh giá phức tạp hơn. Trong giới hạn đồ án này, nhóm chúng em sử dụng 3 hàm heuristic cơ bản và được cài đặt trong file heuristic.py:

H1/ Euclidean distance: khoảng cách “đường chim bay” giữa 2 điểm a, b.

h(a,b) = sqrt((a1-b1)^2 + (a2-b2)^2)

H2/ Manhattan distance: hay còn gọi là “khoảng cách L1” giữa 2 điểm a, b.

h(a,b) = abs(a1-b1) + abs(a2-b2)

H3/ Chebyshev distance: còn gọi là “khoảng cách bàn cờ vua” giữa 2 điểm a, b.

h(a,b) = max(abs(a1-b1), abs(a2-b2))

Mỗi hàm heuristic đều có điểm mạnh và điểm yếu trong từng trường hợp cụ thể khác nhau sẽ được trình bày ở phần kết quả thực nghiệm của báo cáo. Với mỗi loại thuật toán cần dùng hàm heuristic đều sẽ được chạy với cả 3 hàm heuristic này.

# Kết quả thử nghiệm và nhận xét

## Hướng dẫn sử dụng chương trình

Ta chỉ cần làm việc với file main.py, nhận các argument làm tham số.

Giả sử ta đang đứng ở route, có cách tổ chức như sau:

|  |
| --- |
| route/  |--- input/  |--- source/  |--- run.sh |

Nếu muốn, ta chỉ cần đổi thư mục input theo đúng format của giáo viên và cũng được trình bày ở phần bên dưới.

Sau đó đơn giản chạy lệnh:

|  |
| --- |
| bash run.sh  hoặc  sh run.sh |

Chương trình sẽ tự động cài đặt các thư viện cần thiết cũng như chạy các thuật toán theo quy định, các kết quả (mỗi kết quả tương ứng gồm: 1 file .mp4 thể hiện đường đi, 1 file .jpg lưu kết quả cuối cùng, 1 file .txt chứa chi phí đường đi tốt nhất, đối với **level\_3**, các dòng tiếp theosẽ **bổ sung thêm** chi phí tốt nhất thay đổi sau 1000 generation) sẽ lưu trong folder ouput như sau:

|  |
| --- |
| route/  |--- input/  |--- output/  |--- source/  |--- run.sh |

Trong run.sh được cài đặt sẵn câu lệnh:

|  |
| --- |
| python source/main.py |

Ngoài ra cũng có các ví dụ có sẵn để dung các tham số khác.

File main.py có các tham số:

|  |
| --- |
| --input-folder: đường dẫn đến input folder  --output-folder: đường dẫn đến output folder  --print-search-best\_cost: có in ra thêm search best\_cost không trong file txt (mặc định chỉ in best path best\_cost)  --dont-show-video: có hiện video đường đi lên màn hình không (mặc định là có hiện lên)  --auto-close: nếu có hiện đường đi lên, thì đây là tham số số giây tự động tắt giữa các màn hình show lên.  --level: các level thực hiện, mặc định là 1, 2, 3, 4 (advance)  --path: chứa đường dẫn đến 1 file bản đồ (.txt), thay gì thực hiện tất cả các file trong thư mục input, chỉ thực hiện file này. |

Cụ thể có thể coi cài đặt trong file main.py hoặc chạy: python source/main.py -h.

***Lưu ý 1:***

Do sự không ổn định của thư viện convert sang video, có thể trong quá trình chạy: python source/main.py có thể bị dừng giữa chừng (nhưng không báo lỗi) và không chạy hết các file trong folder input.

Hi vọng giáo viên chấm bài có thể chạy lại bằng command:

|  |
| --- |
| python source/main.py --dont-show-video |

để quá trình chạy chương trình ổn định hơn, xin cảm ơn.

Mặc định vẫn show video đường đi lên do chúng em cũng tốn kha khá thời gian để code tính năng đó, cũng như sẽ đem lại hiệu ứng hình ảnh tốt hơn.

Nếu chương trình xảy ra lỗi ***pygame.error: No available video device*** là do không có thiết bị xuất, chúng ta chỉ cần đổi thành câu lệnh:

|  |
| --- |
| python source/main.py --dont-show-video |

***Lưu ý 2:***

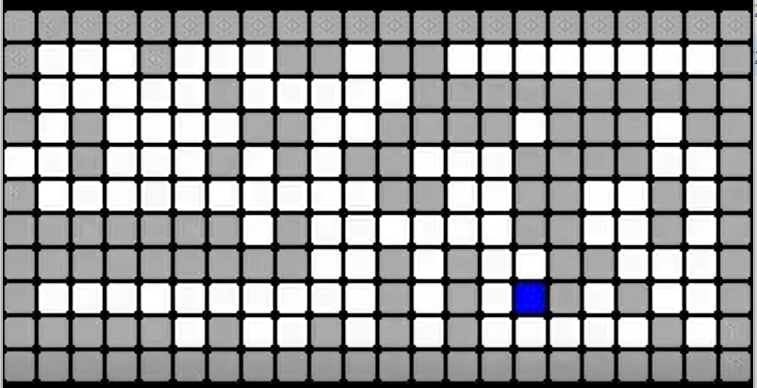
Nhóm chúng em đã thể hiện đường đi trong lúc chạy, cũng như lưu kết quả đường đi dưới dạng file video mp4 cũng như hình ảnh khi kết thúc đường đi với file jpg. Để đảm bảo thời gian chạy đủ nhanh và hiệu quả, chúng em sử dụng kĩ thuật skip frame cũng như tăng tốc độ chạy của video, sử dụng độ phân giải thấp hơn; kết quả lưu được ở video có thể hơi nhanh và kích thước màn hình hiển thị có thể hơi nhỏ. Nhằm đảm bảo tốc độ chạy nên mong thầy cô thông cảm.

## Thử nghiệm

### Bản đồ không có điểm thưởng:

|  |
| --- |
| *Ghi chú:*   * *Ô màu xanh là ô xuất phát* * *Ô màu trắng trên tường là đích đến* * *best\_cost: Chi phí đường đi từ nguồn đến đích, best\_cost = NO nếu không tồn tại đường đi* * *search\_cost: Số lượng các ô được duyệt trong bản đồ* * *Greedy BFS Heuristic 1, A\* Heuristic 1: Hàm Heuristic đánh giá theo* ***euclidean\_distance*** * *Greedy BFS Heuristic 2, A\* Heuristic 2: Hàm Heuristic đánh giá theo* ***manhattan\_distance*** * *Greedy BFS Heuristic 3, A\* Heuristic 3: Hàm Heuristic đánh giá theo* ***chebyshev\_distance*** |

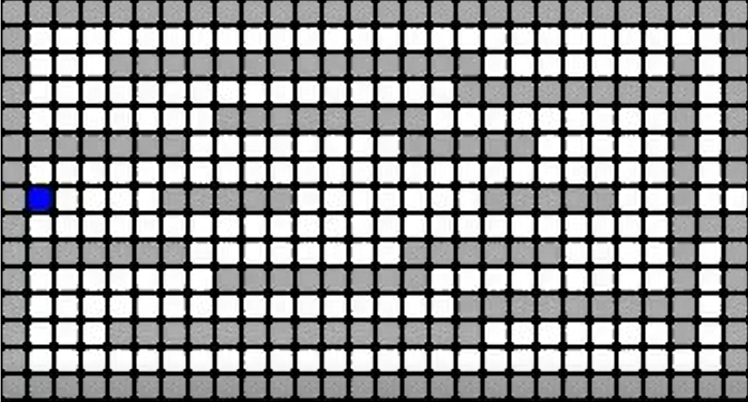
**Bản đồ 1:**

****

|  |  |  |
| --- | --- | --- |
| Hình . DFS (best\_cost = 19, search\_cost = 45) | Hình . BFS (best\_cost = 19, search\_cost = 80) | Hình . UCS (best\_cost = 19, search\_cost = 78) |
| Hình . Greedy BFS Heuristic 1 (best\_cost = 19, search\_cost = 29) | Hình . Greedy BFS Heuristic 2 (best\_cost = 25, search\_cost =29) | Hình . Greedy BFS Heuristic 3 (best\_cost = 19, search\_cost = 19) |
| Hình . A\* Heuristic 1 (best\_cost = 19, search\_cost = 29) | Hình . A\* Heuristic 2 (best\_cost = 19, search\_cost = 29) | Hình . A\* Heuristic 3 (best\_cost = 19, search\_cost = 29) |

***Nhận xét:***

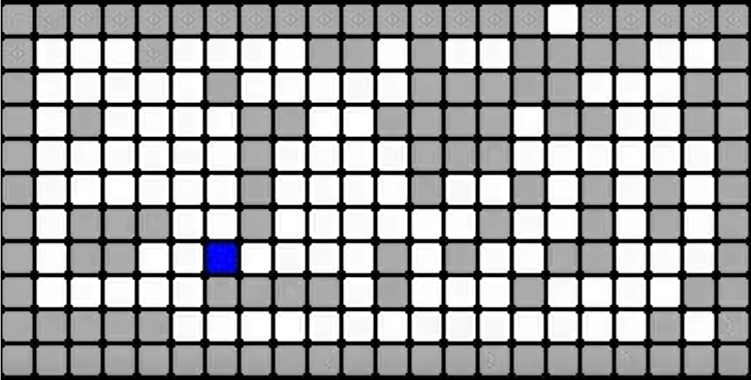
* Tính đầy đủ: Tất cả các giải thuật đều tìm được đường đi và đường đi chưa có sự khác biệt nhiều.
  + Tất cả giải thuật đều có chung tính chất sẽ tìm được đường đi từ nguồn đến đích, nếu tồn tại.
* Tính tối ưu: Nhìn chung, khi xét trên bản đồ này, các đường đi đều khá giống nhau và tối ưu về độ dài, ngoại trừ thuật toán Greedy Best First Search Heuristic 2 (manhattan\_distance) có đường đi dài nhất so với các thuật toán còn lại.
* Độ mở các nút: Đạt kết quả tối nhất là độ mở các nút của thuật toán Greedy Best First Search, hầu như tìm được đường đi trong lần duyệt đầu tiên. Đặc biệt, Greedy BFS Heuristic 3 có search\_cost = best\_cost rơi vào trường hợp tốt nhất và tối ưu nhất. Ngược lại, khi xét trên 2 giải thuật BFS và UCS các nút mở tương đối nhiều gần như là tất cả với BFS có search\_cost = 80, đuổi theo sau là UCS với giá trị 78, điều này thể hiện rõ việc hao tốn chi phí về bộ nhớ và thời gian chạy cho việc này. Riêng A\* có độ mở tương đối ổn định và chiếm một phần nhỏ với cùng giá trị search\_cost = 29 trên cả 3 hàm Heuristic khác nhau.
* Chi phí đường đi: Trừ trường hợp thuật toán Greedy Best First Search Heuristic 2 có chi phí lớn nhất best\_cost = 25 thì các giải thuật còn lại đều có cùng giá trị best\_cost = 19.
* **Bản đồ 2:**

****

|  |  |  |
| --- | --- | --- |
| Hình . DFS (best\_cost = 94, search\_cost =228 ) | Hình . BFS (best\_cost = 46, search\_cost =233) | Hình . UCS (best\_cost = 46, search\_cost =233) |
| Hình . Greedy BFS Heuristic 1 (best\_cost = 62, search\_cost =173) | Hình . Greedy BFS Heuristic 2 (best\_cost = 60, search\_cost =181) | Hình . Greedy BFS Heuristic 3 (best\_cost = 82, search\_cost =168) |
| Hình . A\* Heuristic 1 (best\_cost = 46, search\_cost = 233) | Hình . A\* Heuristic 2 (best\_cost = 46, search\_cost = 233) | Hình . A\* Heuristic 3 (best\_cost = 46, search\_cost = 233) |

***Nhận xét:***

* Tính đầy đủ: Tất cả các giải thuật đều tìm được đường đi.
* Tính tối ưu: Trong bản đồ 2, các giải thuật dường như đã tỏ rõ sự khác biệt của chúng. Nếu DFS loang theo chiều sâu, phát triển một nhánh để đến được đích nên đường đi tìm được khá dài, kém hiệu quả. Ngược lại với DFS, BFS và UCS lại tìm kiếm loang theo chiều rộng nên trong trường hợp này có phần hiệu quả và tìm được tìm đi tối ưu nhất.
  + Sở dĩ, UCS tìm được đường đi khá giống với BFS trong các bản đồ là vì UCS tìm kiếm theo trọng số (chi phí) thấp của các ô kề nó, mà trong phạm vi đồ án ta xét tất cả các ô đều có trọng số bằng 1, nghĩa là chi phí bằng nhau ở mọi hướng. Mặt khác, BFS và UCS đều đảm bảo tìm được đường đi ngắn nhất nên vấn đề khá giống nhau về đường đi có thể giải thích được.
* Độ mở các nút: Các giải thuật có sự khác biệt
  + Trường hợp xấu nhất: BFS, UCS, A\* Heuristic 1 phải duyệt toàn bộ các nút để tìm được đường đi tối ưu.
  + Trường hợp tương đối xấu: DFS gần như duyệt toàn bộ các nút
  + Trường hợp còn lại: Ngoại trừ các thuật toán nêu trên, các thuật toán còn lại duyệt tương đối và hiệu quả hơn.
* Chi phí đường đi: Dựa trên từng tính chất riêng của từng loại thuật toán mà chi phí về đường đi của mỗi giải thuật có sự khác biệt:
  + Chi phí thấp nhất (best\_cost = 46): BFS, UCS và các giải thuật A\*. Nghĩa là, các giải thuật này đảm bảo rằng nếu tồn tại đường đi từ nguồn đến đích thì đường đi tìm được là đường đi ngắn nhất, có chi phí thấp nhất.
* **Bản đồ 3:**

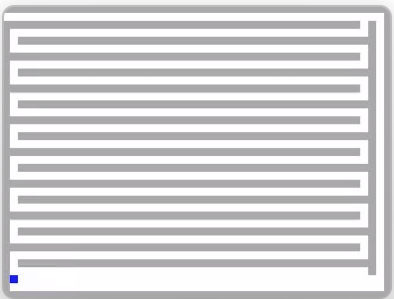


|  |  |  |
| --- | --- | --- |
| Hình . DFS (best\_cost = NO, search\_cost = 117) | Hình . BFS (best\_cost = NO, search\_cost = 117) | Hình . UCS (best\_cost = NO, search\_cost = 117) |
| Hình . Greedy BFS Heuristic 1 (best\_cost = NO, search\_cost = 117) | Hình . Greedy BFS Heuristic 2 (best\_cost = NO, search\_cost = 117) | Hình . Greedy BFS Heuristic 3 (best\_cost = NO, search\_cost = 117) |
| Hình . A\* Heuristic 1 (best\_cost = NO, search\_cost = 117) | Hình . A\* Heuristic 2 (best\_cost = NO, search\_cost = 117) | Hình . A\* Heuristic 3 (best\_cost = NO, search\_cost = 117) |

***Nhận xét:***

Bản đồ trên là trường hợp không tồn tại đường đi giữa nguồn và đích nên kết quả thu được của mọi giải thuật là như nhau:

* Tính đầy đủ: Tất cả các giải thuật đều không tìm được đường đi.
* Tính tối ưu: Các giải thuật rơi vào trường hợp xấu nhất nên không có giải thuật nào tối ưu.
* Độ mở các nút: Trừ những nút không thể duyệt đến thì toàn bộ nút còn lại đều được duyệt.
* Chi phí đường đi: best\_cost = NO không tồn tại đường đi.
* **Bản đồ 4:**



|  |  |  |
| --- | --- | --- |
| Hình . DFS (best\_cost = 126, search\_cost = 173) | Hình . BFS (best\_cost = 126, search\_cost = 390) | Hình . UCS (best\_cost = 126, search\_cost = 386) |
| Hình . Greedy BFS Heuristic 1 (best\_cost = 738, search\_cost = 863) | Hình . Greedy BFS Heuristic 2 (best\_cost = 738, search\_cost = 859) | Hình . Greedy BFS Heuristic 3 (best\_cost = 738, search\_cost = 868) |
| Hình . A\* Heuristic 1 (best\_cost = 316, search\_cost = 316) | Hình . A\* Heuristic 2 (best\_cost = 126, search\_cost = 312) | Hình . A\* Heuristic 3 (best\_cost = 126, search\_cost = 316) |

***Nhận xét:***

* Tính đầy đủ: Tất cả các giải thuật đều tìm được đường đi.
* Tính tối ưu: Bản đồ 4 xét về tính tối ưu chia rõ giữa:
  + Tối ưu: DFS, BFS, UCS, A\* Heuristic
  + Không tối ưu: Greedy Best First Search
    - Gready BFS chỉ quan tâm đến h(x) – khoảng cách heuristic từ vị trí hiện tại đến đích nên trong trường hợp bản đồ này, thuật toán đã bị “lừa”. Thuật toán chỉ cố gắng đi đến các điểm có vẻ gần với địch mà không nhận ra đó chính là đường dài, “vòng vèo” hơn rất nhiều so với đường tắt ngắn hơn nhưng chi phí để đi đến lúc đầu có vẻ nhiều hơn.
* Độ mở các nút: Trường hợp tốt nhất là DFS với số lượng nút được duyệt đến là ít nhất. Nhưng điều này không “ổn định”, phụ thuộc vào thứ tự nút ta chọn để duyệt, và may mắn trong trường hợp này thứ tự nút duyệt phù hợp để DFS có thể thuận lợi tìm ra đường tốt nhất. Kế đến là các giải thuật A\*, BFS, UCS tương đối tốt. Ngược lại, Greedy Best First Search khá xấu gần như mở tất cả nút của bản đồ.
* Chi phí đường đi:
  + Tối ưu (best\_cost = 126): DFS, BFS, UCS, A\* Heuristic
  + Không tối ưu: Greedy Best First Search
* **Bản đồ 5:**



|  |  |  |
| --- | --- | --- |
| Hình . DFS (best\_cost = 247, search\_cost = 296) | Hình . BFS (best\_cost = 83, search\_cost = 1374) | Hình . UCS (best\_cost = 83, search\_cost = 1370) |
| Hình . Greedy BFS Heuristic 1 (best\_cost = 83, search\_cost = 151) | Hình . Greedy BFS Heuristic 2 (best\_cost = 105, search\_cost = 160) | Hình . Greedy BFS Heuristic 3 (best\_cost = 83, search\_cost = 151) |
| Hình . A\* Heuristic 1 (best\_cost = 83, search\_cost = 788) | Hình . A\* Heuristic 2 (best\_cost = 83, search\_cost = 685) | Hình . A\* Heuristic 3 (best\_cost = 83, search\_cost = 843) |

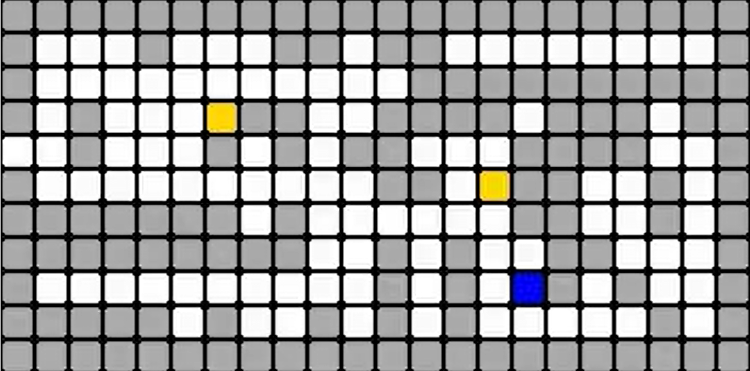
***Nhận xét:***

* Tính đầy đủ: Tất cả các giải thuật đều tìm được đường đi.
* Tính tối ưu: Các giải thuật dường như đều tìm được đường đi tối ưu nhất, trừ DFS và Greedy BFS 2. Trong số đó, DFS kém tối ưu nhất, nguyên nhân thì như đã đề cập ở, do việc chọn thứ tự đỉnh quyết định rất nhiều đến kết quả của DFS. Do trong việc cài đặt DFS nhóm chúng em chọn ưu tiên đi theo hướng trái phải rồi mới di chuyển lên xuống và thuật toán sẽ kết thúc ngay khi đến được đích nên mới sinh ra kết quả không tối ưu của DFS trong trường hợp này. Còn về Greedy BFS 2 (sử dụng độ đo manhattan) cũng không tìm được đường đi tối ưu.
* Độ mở các nút: Đổi lại từ chi phí “đắt đỏ” về đường đi của DFS thì thuật toán này đã có kết quả độ mở hết sức hiệu quả với số lượng các nút được mở khá ít ỏi. Ngược lại, BFS và UCS thì rơi vào trường hợp dường như duyệt đến tất cả các nút có trong bản đồ nên quá tốn kém về mặt thời gian cũng như bộ nhớ. Tiếp đến, với số lượng nút tương đối được duyệt là giải thuật A\* chiếm một nửa số nút. Còn lại các giải thuật Greedy BFS cũng khá hiệu quả.
* Chi phí đường đi: Gần như toàn bộ giải thuật đều tìm được đường đi có chi phí thấp nhất với best\_cost = 83. Ngược lại, DFS có chi phí cao nhất best\_cost = 247. Ngoài ra, Greedy BFS 2 cũng có chi phí đường đi không được tối ưu với best\_cost = 105

### Bản đồ có điểm thưởng

|  |
| --- |
| *Ghi chú:*   * *Ô màu vàng: Ô điểm thưởng* * *Các lưu ý khác như bản đồ không có điểm thưởng* |

* **Trường hợp bản đồ có 2 điểm thưởng:**
  + **Bản đồ 1:**

****

|  |  |  |
| --- | --- | --- |
| Hình . A\* Heuristic 1 (best\_cost = 16, search\_cost = 80) | Hình . A\* Heuristic 2 (best\_cost = 16 , search\_cost = 79) | Hình . A\* Heuristic 3 (best\_cost = 16, search\_cost = 81) |

***Nhận xét:***

Với sự xuất hiện của điểm thưởng, trong trường hợp này bản đồ có 2 điểm thưởng nhưng ta có thể quan sát thấy chỉ 1 điểm thưởng được duyệt đến và đi qua, trong khi điểm thưởng thứ 2 chưa được xét đến. Bản đồ 1 này vẫn chưa thấy rõ được sự khác biệt, cụ thể:

* Tính đầy đủ: Tất cả các giải thuật đều tìm được đường đi.
* Tính tối ưu: Tất cả giải thuật đều tìm được đường đi tối ưu nhất theo tính chất của A\* đã được trình bày.
* Độ mở các nút: Có sự chênh lệch khá nhỏ về số lượng nút được duyệt trong bản đồ này.
* Chi phí đường đi: Tất cả các giải thuật đều tìm được đường đi có chi phí đường đi thấp nhất best\_cost = 16.
* **Trường hợp bản đồ có 5 điểm thưởng:**
  + **Bản đồ 2:**

Ảnh có chứa bàn

Mô tả được tạo tự động

|  |  |  |
| --- | --- | --- |
| Hình . A\* Heuristic 1 (best\_cost = -57, search\_cost = 122) | Hình . A\* Heuristic 2 (best\_cost = -57, search\_cost = 122) | Hình . A\* Heuristic 3 (best\_cost = -57, search\_cost = 122) |

***Nhận xét:***

* Tính đầy đủ: Tất cả các giải thuật đều tìm được đường đi giống nhau, và với số lượng điểm thưởng mà cả 3 giải thuật mang về được là 3/5.
* Tính tối ưu: Tất cả giải thuật đều tìm được đường đi tối ưu và thật bất ngờ tuy với 3 hàm Heuristic khác nhau nhưng đường đi tìm đường thì không có gì khác biệt.
* Độ mở các nút: Rõ ràng ngoại trừ những ô không thể duyệt đến thì toàn bộ các nút đều được duyệt. Ở đây, search\_cost có giá trị là
* Chi phí đường đi: Tất cả các giải thuật đều tìm được đường đi có chi phí đường đi khá tốt best\_cost = -57.
* **Trường hợp bản đồ có 10 điểm thưởng:**
  + **Bản đồ 5:**

****

|  |  |  |
| --- | --- | --- |
| Hình . A\* Heuristic 1 (best\_cost = 71, search\_cost = 826) | Hình . A\* Heuristic 2 (best\_cost = 71, search\_cost = 828) | Hình . A\* Heuristic 3 (best\_cost = 71, search\_cost = 851) |

***Nhận xét:***

Đường đi tìm được tối ưu với chi phí đường đi best\_cost = 71 và 4/10 điểm thưởng được duyệt cho cả 3 trường hợp. Tuy nhiên, với sự khác biệt về tính chất của từng hàm Heuristic cho nên cả 3 đều có đường đi khác nhau dù cùng một chi phí. Độ mở các nút cũng khá tương đối nhiều, nhưng một số điểm thưởng vẫn chưa được duyệt đến, nếu trong trường hợp các điểm thưởng chưa duyệt giảm được chi phí khá lớn thì rõ ràng rằng đường đi tìm được chưa hẳn là tối ưu nhất. Qua ví dụ này, ta cũng kết luận được rằng giải thuật A\*chưa chắc sẽ tìm được đường đi ngắn nhất và tùy thuộc khá lớn vào hàm Heuristic.

### Bản đồ có điểm đón

* **Trường hợp bản đồ có 5 điểm đón:**
  + **Bản đồ 1:**

|  |  |
| --- | --- |
|  | Ảnh có chứa bàn  Mô tả được tạo tự động  Hình . UCS + Genetic Algorithm (best\_cost = search\_cost = 63) |

***Nhận xét:***

Với số lượng điểm đón nhỏ là 5, số hoán vị các điểm đón có thể là 5! = 120, do đó với số lượng phần tử trong quần thể (~50), số lượng vòng đời lớn (20000) thì thuật toán có thể đảm bảo tìm đường đi tối ưu, tuy nhiên do phụ thuộc lớn vào ngẫu nhiên nên tuy thử nghiệm chưa gặp nhưng về đánh giá và tính toán, thuật toán vẫn có xác suất cho kết quả không tối ưu:

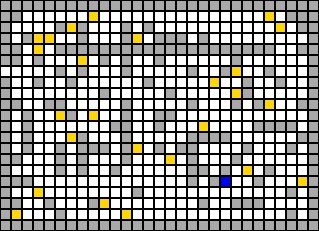
* Tính đầy đủ: Giải thuật tìm được đường đi.
* Tính đa định: Mỗi lần thực thi, thuật toán thực hiện khác nhau do phụ thuộc vào quá trình ngẫu nhiên.
* Tính tối ưu: Giải thuật đều tìm được đường đi tối trong các lần thử nghiệm, tuy nhiên tính ổn định không được đảm bảo do phụ thuộc yếu tố ngẫu nhiên.
* Chi phí đường đi: Giải thuật tìm được đường đi có chi phí đường đi thấp nhất là 63 (đã thử nghiệm với brute-force) ở các lần thử nghiệm.
* **Trường hợp bản đồ có 10 điểm đón:**
  + **Bản đồ 2:**

|  |  |
| --- | --- |
|  | Hình . UCS + Genetic Algorithm (best\_cost = search\_cost = 50) |

***Nhận xét:***

Với số lượng điểm đón ở mức vừa phải là 10, số hoán vị các điểm đón có thể là 10! = 3628800 thì số lần thuật toán tìm được đường đi tối ưu là hơn 70% theo số lần thống kê của nhóm, tuy nhiên, vẫn có trường hợp, thuật toán rơi vào cục bộ địa phương nên chỉ tìm được đường đi gần tối ưu.

* Tính đầy đủ: Giải thuật tìm được đường đi.
* Tính đa định: Mỗi lần thực thi, thuật toán thực hiện khác nhau do phụ thuộc vào quá trình ngẫu nhiên.
* Tính tối ưu: Giải thuật tìm được đường đi tối ưu trong hơn 70% lần thử nghiệm. Các trường hợp còn lại là gần tối ưu.
* Chi phí đường đi: Giải thuật tìm được đường đi có chi phí đường đi thấp nhất là 50 (đã thử nghiệm với brute-force) và kết quả gần tối ưu là 52.
* **Trường hợp bản đồ có 25 điểm đón:**
  + **Bản đồ 3:**



|  |  |  |
| --- | --- | --- |
| Ảnh có chứa tòa nhà  Mô tả được tạo tự động  Hình . UCS + Genetic Algorithm (best\_cost = search\_cost = 184) | Ảnh có chứa tòa nhà  Mô tả được tạo tự động  Hình . UCS + Genetic Algorithm (best\_cost = search\_cost = 162) | Hình . UCS + Genetic Algorithm (best\_cost = search\_cost = 150) |

***Nhận xét:***

Với số lượng điểm đón ở mức lớn, số hoán vị các điểm đón có thể là 25!≈1.5x1025 thì việc tìm ra lời giải tối ưu là không đảm bảo. Vì thế, Genetic Algorithm giúp tìm kết quả tối ưu cục bộ trong thời gian hạn định. Ở mỗi lần thực thi, tùy vào việc ngẫu nhiên mà tìm được lời giải tối ưu nhất có thể. Cụ thể trong 3 thử nghiệm trên là 3 kết quả đạt được tương ứng với lời giải khác nhau phụ thuộc vào cực tiếu địa phương mà “quần thể” của thuật toán đang tập trung.

* Tính đầy đủ: Giải thuật tìm được đường đi.
* Tính đa định: Mỗi lần thực thi, thuật toán thực hiện khác nhau do phụ thuộc vào quá trình ngẫu nhiên.
* Tính tối ưu: Giải thuật không đảm bảo tìm được đường đi tối ưu, mà là cực tiểu địa phương (gần tối ưu), tuy nhiên về mặt toán học, vẫn có khả năng thuật toán tìm kết quả tốt nhất.
* Chi phí đường đi: Giải thuật tìm được đường đi có chi phí thấp nhất trong các lần thử nghiệm là 150, ngoài ra các kết quả khác cũng gần với mức này như 162 và 184...

# Cách tổ chức chương trình và ghi chú lập trình

# *Thư mục bài nộp bao gồm:*

– File run.sh: file shell script chứa các câu lệnh để chạy toàn bộ các thuật toán cho tất cả các bản đồ.

– File report.pdf: báo cáo trình bày các vấn đề được yêu cầu trong đề bài.

– Folder source bao gồm:

+ Thư mục visual:

+ cell.py: chứa các thông tin hỗ trợ việc vẽ 1 cell.

+ map.py: có các hàm chứa thông tin của 1 map, các hàm hỗ trợ liên quan đến map cũng như việc thực hiện vẽ cũng như show map.

+ util.py: chứa các hàm hỗ trợ thường dùng.

+ path.py: chứa template của 1 class “thuật toán”, quy định các hàm chung nhất của các thuật toán tìm đường đi.

+ algo.py: chứa cài đặt của 5 thuật toán cơ bản.

+ award\_algo.py: chứa cài đặt của thuật toán cho bản đồ điểm thưởng và bản đồ điểm đón.

+ tele\_algo.py: chứa cài đặt của thuật toán cho bản đồ có điểm dịch chuyển.

+ heuristic.py: chứa cài đặt 3 hàm heuristic cơ bản.

+ main.py: chứa cài đặt tổng, cũng như là xử lý các ArgumentParser.

+ requirements.txt: chứa các thư viện cần cài thông qua pip.

– Folder input bao gồm các folder, file theo format và yêu cầu của giáo viên:

+ level\_1: folder chứa 5 map (.txt) của level\_1.

+ level\_2: folder chứa 3 map (.txt) của level\_2.

+ level\_3: folder chứa 5 map (.txt) của level\_3.

+ advance: folder chứa 4 map (.txt) của level advance.

# *Mô tả file run.sh*

- Thực hiện cài đặt các thư viện cần thiết thông qua pip từ requirements.txt.

- Chạy file main.py.

- Có các ví dụ sử dụng các argument với main.py.

- Cụ thể các argument đã được trình bày ở mục Hướng dẫn sử dụng chương trình.

# Tài liệu tham khảo

[1] Các slide về giải thuật và tài liệu hướng dẫn được cung cấp trên Moodle.

[2] Juwairiah, Juwairiah & Pratama, Dicky & Rustamaji, Heru & Sofyan, Herry & Prasetyo, Dessyanto. (2019). Genetic Algorithm for Optimizing Traveling Salesman Problems with Time Windows (TSP-TW). International Journal of Artificial Intelligence & Robotics (IJAIR). 1. 1. 10.25139/ijair.v1i1.2024.

[3] https://vi.wikipedia.org/wiki/Bài\_toán\_đường\_đi\_ngắn\_nhất

[4] <https://algs4.cs.princeton.edu/44sp/>

[5] <https://viblo.asia/>

[6] Slide về môn Trí tuệ nhân tạo của thầy Phạm Minh Tuấn