

Bài tập nhóm cuối kì:
Bài toán vận chuyển gói hàng theo thời gian thực với đa tác tử

Mô tả

Cho một bảng A có kích thước $N \times N$, trong đó giá trị $A_{i,j} = 0$ ứng với ô trống và $A_{i,j} = 1$ ứng với ô có vật cản (không thể di chuyển vào) với i là dòng, j là cột, $1 \leq i, j \leq N$.

Tại thời điểm $t = 0$, có C robot làm nhiệm vụ vận chuyển hàng hoá có tọa độ (x_i, y_i) với $0 \leq i < C$. Mỗi robot đứng ở một ô không phải vật cản và không có hai robot nào đứng cùng đứng một ô.

Tại mỗi thời điểm $0 < t < T$ (T là thời gian lớn nhất), mỗi robot có thể thực hiện một trong 5 hành động di chuyển: S, L, R, U, D và nhặt gói hàng 1/hoặc trả gói hàng: 0. Ứng với các hành động sau:

S: Đứng tại chỗ.

L: Di chuyển sang trái.

R: Di chuyển sang phải.

U: Di chuyển lên trên.

D: Di chuyển xuống dưới.

1: Nhặt hàng, nếu ô đang đứng, sau khi di chuyển có gói hàng.

0: Không làm gì

2: Thả gói hàng, nếu ô đang đứng, sau khi di chuyển là ô mục tiêu của gói hàng.

Tại mỗi thời điểm $0 \leq t < T$, có một số yêu cầu chuyển các gói hàng sẽ xuất hiện. Mỗi gói hàng sẽ có thông tin: $g^i = \langle s_x^i, s_y^i, e_x^i, e_y^i, e_t^i \rangle$ trong đó với $i \ll G$

s_x^i, s_y^i : Tọa độ hàng và cột của điểm bắt đầu của gói hàng thứ i

e_x^i, e_y^i, e_t^i : Tọa độ hàng và cột của điểm kết thúc, và hạn giao hàng. Quá thời hạn này sẽ bị tính là quá hạn.

Mỗi gói hàng được vận chuyển đúng hạn sẽ được tính tiền công vận chuyển là $r \geq 0$, quá hạn là $r_p \geq 0$. Trong thiết lập này, giá trị phần thưởng được chọn cố định là $r = 10$ và $r_p = 1$.

Mỗi robot có một chi phí di chuyển cố định là $r_c = -0.01$.

Chi phí này chỉ tính khi robot di chuyển (thực hiện hành động LRUD).

Yêu cầu:

Tối ưu phần thưởng nhận được sau T timestep. Cụ thể, gọi phần thưởng nhận được tại thời gian t của robot i là: r_t^i

$$\max \sum_t^i r_t^i \text{ với } 0 \leq i < C, 0 < t \leq T$$

Ví dụ:

$t = 0$:

Chúng ta có bản đồ 5x5:

1 1 1 1 1

1 0 0 0 1

1 0 0 0 1

1 0 0 0 1

1 1 1 1 1

Có 2 robot ở vị trí: (1, 2) (2, 2)

Có 2 gói hàng ở vị trí: $g^1 = (1, 1, 2, 3, 3, 10)$, $g^2 = (2, 2, 3, 3, 3, 10)$

Tại thời điểm $t = 1$:

Thực hiện hành động (S1, R1)

Chi phí trả về là $r_1 = -0.02$ và 2 robot đã nhận hai gói hàng.

Tiếp tục chúng ta có

$t = 2$: (R0, D0), với $r_2 = -0.02$

$t = 3: (R0, D2)$, với $r_3 = 9.98$

$t = 4: (D0, L0)$, với $r_4 = -0.02$

$t = 5: (D2, S0)$, với $r_5 = 10.00$

Ta được tổng cộng phần thưởng là: 10.92

Một số lưu ý khi di chuyển:

- Robot có va chạm, nếu hai robot đứng cùng 1 ô sau khi di chuyển thì robot có chỉ số bé hơn sẽ được ưu tiên hơn.
- Các robot sẽ thực hiện di chuyển trước, sau đó sẽ đến các thao tác với gói hàng. Mỗi robot chỉ được mang tối đa 1 gói hàng và khi đến đích thì mới đc trả gói hàng.
- Nếu trong 1 thao tác lấy gói hàng, có nhiều gói hàng tại một vị trí thì gói hàng có chỉ số thấp nhất (xuất hiện sớm nhất) sẽ được ưu tiên nhất.
- Robot không được đi ra ngoài map và đi vào vật cản.

Các thiết lập:

- $T \leq 2000$, $N \leq 20$, $C \leq 20$, $G \leq 1000$.
- Giá trị T , N , C , G map kiểm tra sẽ được biết trước. Có 5 thiết lập khác nhau.
- Robot sẽ xuất hiện ngẫu nhiên trên map tại thời điểm $T=0$.

Cài đặt

Các bạn xem trong code kèm theo. Các bạn cần gửi lại source code cài đặt và notebook chạy được qua hệ thống kaggle/google collab với một số cấu hình cụ thể, sẽ được công bố ở tuần sau. Notebook không được sử dụng internet.

Thời gian chạy, không quá 1 tiếng trên kaggle (không kể thời gian training).

Chấm điểm:

- Báo cáo: 10 điểm
 - + Ghi rõ thành viên và đóng góp các thành viên
 - + Ghi rõ cách tiếp cận, cài đặt, các phương pháp đã thử nghiệm.
 - + So sánh các phương pháp với nhau.

- + Có điểm cộng nếu phương pháp có tính sáng tạo, đạt hiệu quả cao.
- Mức độ tối ưu (so với baseline BFS): 10 điểm
 - + Các bạn sẽ có 1 số config để chạy và tối ưu kết quả.
 - + Các bạn sẽ được điểm cộng nếu phương pháp có thể tối đa được: số hàng giao đúng hạn, khoảng thời gian di chuyển không gói hàng nhỏ nhất, tổng quãng đường di chuyển nhỏ nhất (ngoài các kết quả thông thường)....
 - + Điểm cộng tối đa: 2 điểm
- Code: 5 điểm.
- Trình bày (vấn đáp): 5 điểm

Gợi ý:

- Cách tiếp cận đồ thị: BFS tham lam, hoặc tối ưu đàn kiến
- Cách tiếp cận nâng cao: học sâu tăng cường, học sâu trên đồ thị.