

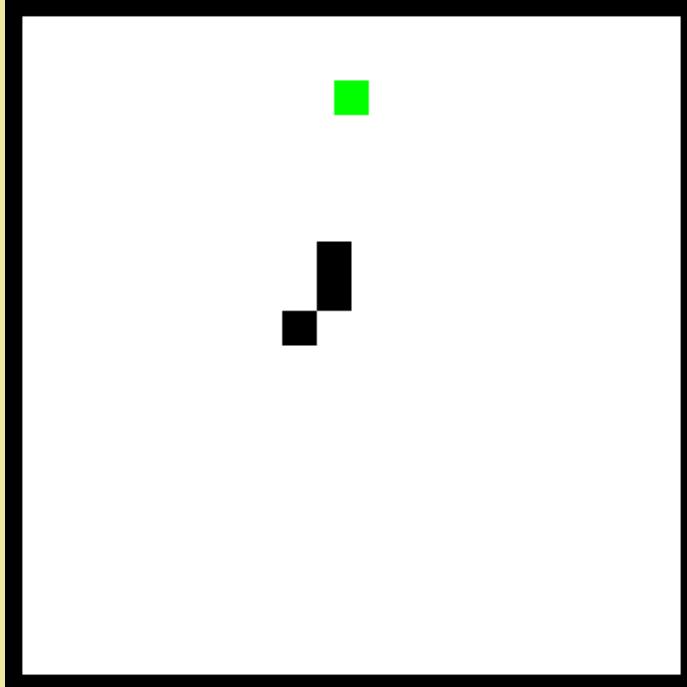


SNAKE GAME AND DRL

**DÙNG DRL ĐỂ CHƠI
RẮN SĂN MỒI**

Le Nhung

CÁC THÀNH PHẦN TRONG GAME



1

Tác nhân (agent): con rắn, thu thập thông tin về trạng thái hiện tại học cách điều hướng môi trường để ăn thức ăn và tránh đâm vào tường

2

Môi trường (environment):

- bao gồm lưới 2D xung quanh rắn, cơ thể rắn, thức ăn và chướng ngại vật (4 phía xung quanh)
- thưởng hoặc phạt tác nhân dựa trên hành động đã thực hiện

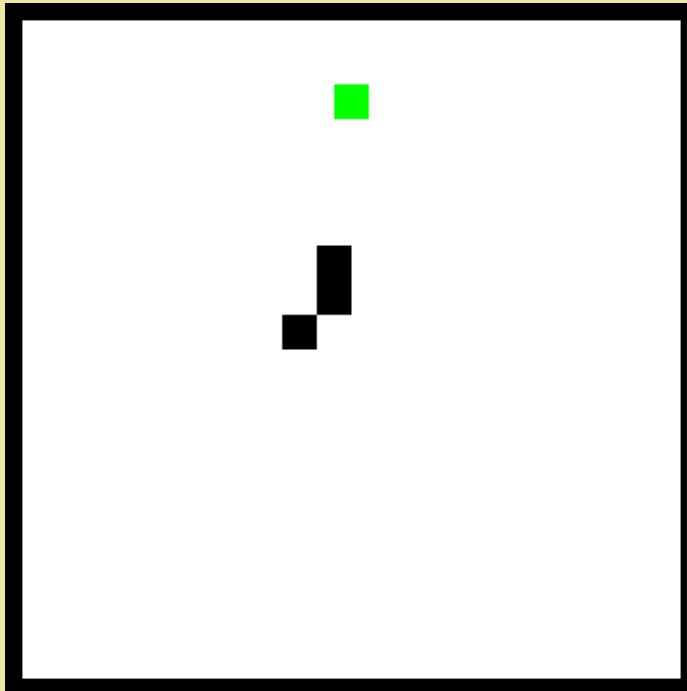
3

Trạng thái (state):

- Nguy hiểm phía trước
- Nguy hiểm bên phải
- Nguy hiểm bên trái
- Đồ ăn bên trái
- đồ ăn bên phải
- Đồ ăn bên trên
- Đi thẳng
- Rẽ phải
- Rẽ trái
- Đi xuống



CÁC THÀNH PHẦN TRONG GAME



4

Hành động (action):

- đi thẳng
- rẽ trái
- rẽ phải

5

Chính sách (policy):

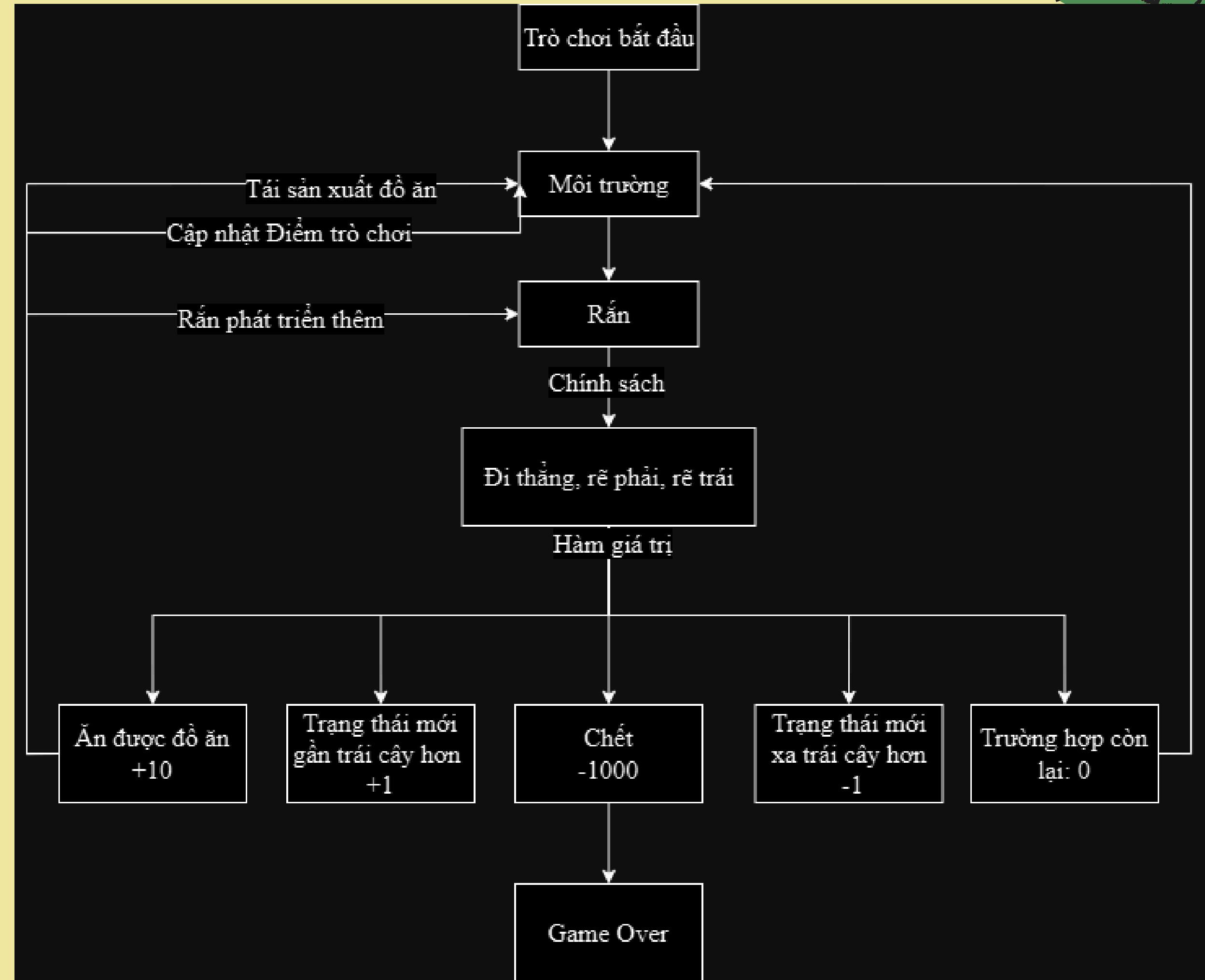
- tập hợp các quy tắc mà tác nhân tuân theo để xác định bước đi tiếp theo dựa trên trạng thái hiện tại
- sử dụng thuật toán Q-learning, điều chỉnh chiến lược của mình dựa trên hệ thống phần thưởng và hình phạt để tìm ra con đường tối ưu dẫn đến thành công

6

Phần thưởng (reward):

- Ăn trái cây: 10
- Trạng thái mới gần trái cây hơn so với trạng thái trước đó: 1
- Trạng thái mới xa trái cây hơn so với trạng thái trước đó: -1
- Chết: -1000
- Bất kỳ phần thưởng nào khác: 0

FLOW CHART



CÁC THÀNH PHẦN CỤ THỂ

Trạng thái tác nhân

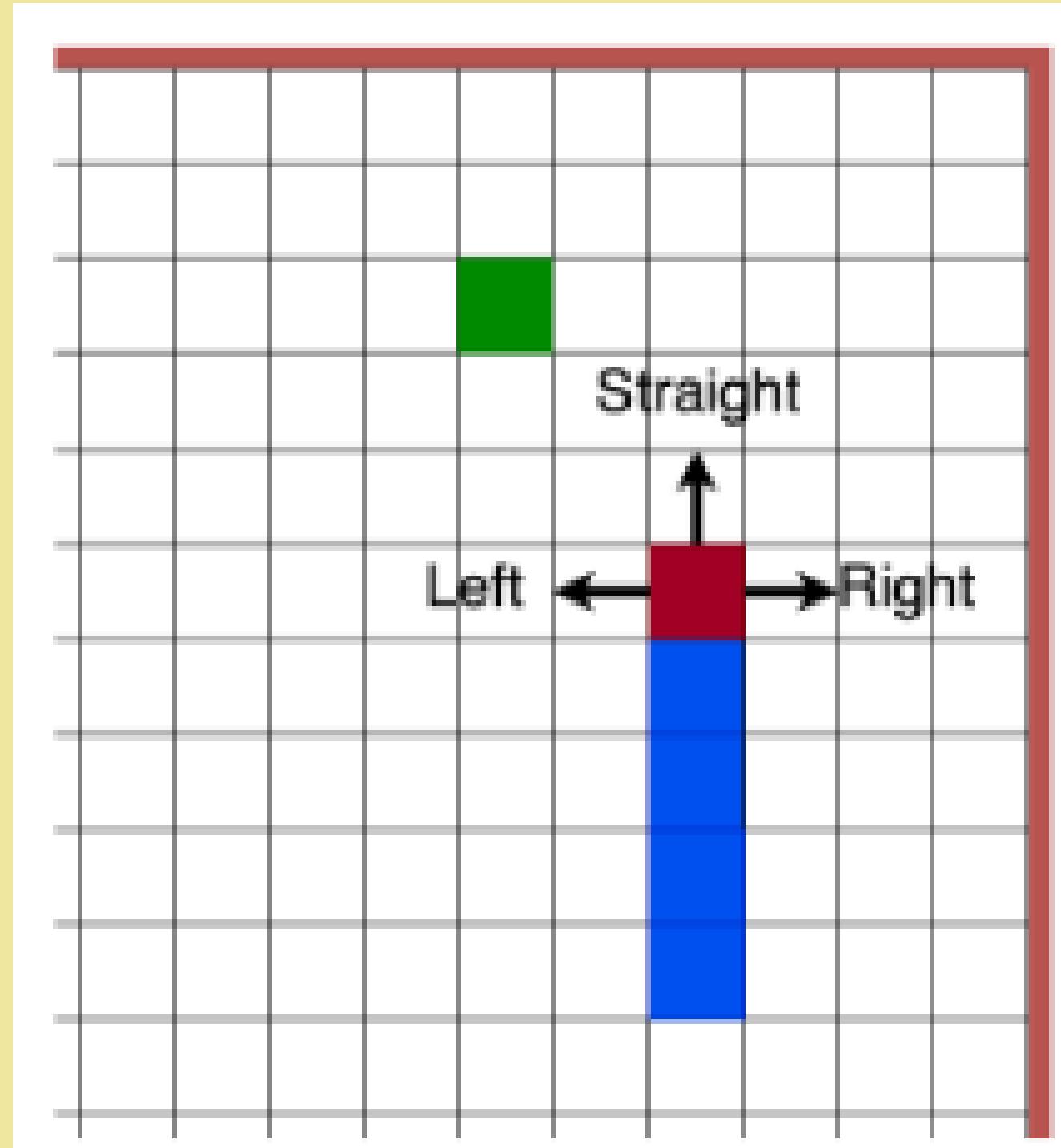
- sự phản ánh của trò chơi rắn mà tác nhân có thể hiểu
- chứa dữ liệu mà tác nhân phải xử lý để quyết định hành động nào cần thực hiện
- ảnh hưởng đến hiệu suất hoạt động của tác nhân và tốc độ xử lý thông tin
- Nếu cung cấp nhiều dữ liệu có liên quan hơn, tác nhân sẽ hoạt động tốt hơn nhiều nhưng xử lý thông tin chậm hơn
- Mỗi giá trị, có thể là 0 hoặc 1, biểu thị một điều khác nhau

No	State	Value
1	Danger straight	0 or 1
2	Danger right	0 or 1
3	Danger left	0 or 1
4	Direction left	0 or 1
5	Direction right	0 or 1
6	Direction up	0 or 1
7	Direction down	0 or 1
8	Food-left	0 or 1
9	Food-right	0 or 1
10	Food-up	0 or 1
11	Food-down	0 or 1

CÁC THÀNH PHẦN CỤ THỂ

Hành động

- những hành động mà tác nhân có thể thực hiện dựa trên trạng thái cụ thể
- Tác nhân chỉ có thể thực hiện ba hành động:
 - rẽ phải
 - rẽ trái
 - đi theo đường thẳng



CÁC THÀNH PHẦN CỤ THỂ

Phần thưởng

Phần thưởng dựa trên hành động mà tác nhân đã làm trước đó

- ăn trái cây: +10
- trạng thái mới gần trái cây hơn trạng thái cũ: +1
- trạng thái mới xa trái cây hơn trạng thái cũ: -1
- chết: -1000
- các trường hợp khác: 0

No	Function	Value
1	Eat food	+10
2	New state near food than old state	+1
3	New state far food than old state	-1
4	Die	-1000
5	Else	0

CÁC BƯỚC CỤ THỂ

Bước 1: Khởi tạo trò chơi

- Khởi tạo khung trò chơi 800 x 600 làm môi trường
- Khởi tạo rắn và thức ăn

Bước 2: Bắt đầu trò chơi

- Quan sát trạng thái: Ghi lại trạng thái hiện tại của trò chơi (vị trí rắn, vị trí thức ăn, v.v.).



CÁC BƯỚC CỤ THỂ

Bước 3: Dùng policy để dự đoán hướng đi tiếp theo

- Sử dụng mạng nơ-ron để đánh giá trạng thái hiện tại và chọn hành động (di chuyển lên, xuống, trái, phải)
- Di chuyển rắn dựa trên hành động đã chọn
- Cập nhật trạng thái trò chơi dựa trên hành động
- Kiểm tra va chạm (rắn với tường hoặc chính nó)



CÁC BƯỚC CỤ THỂ

Bước 4: Tính toán phần thưởng

- Phần thưởng dựa trên hành động mà tác nhân đã làm trước đó
- ăn trái cây: +10
- trạng thái mới gần trái cây hơn trạng thái cũ: +1
- trạng thái mới xa trái cây hơn trạng thái cũ: -1
- chết: -1000
- các trường hợp khác: 0
- Lưu trải nghiệm (trạng thái, hành động, phần thưởng, trạng thái tiếp theo) để học trong tương lai



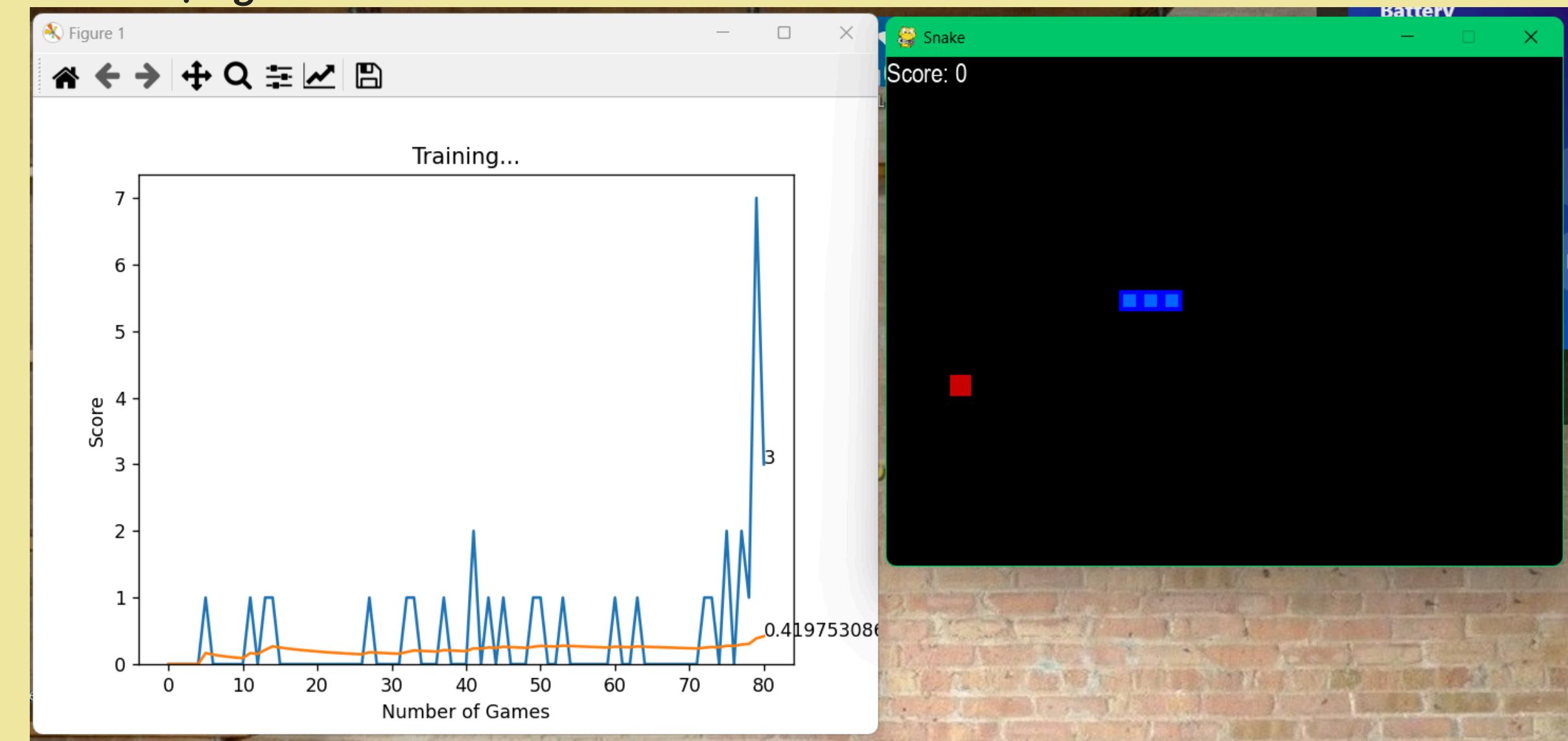
CÁC BƯỚC CỤ THỂ

Bước 5: Đào tạo mạng nơ-ron

- Sử dụng các trải nghiệm đã lưu trữ để đào tạo mạng nơ-ron
- Điều chỉnh giá trị Q dựa trên phần thưởng và hành động đã thực hiện
- Kiểm tra xem trò chơi đã kết thúc chưa (phát hiện va chạm)
- Nếu chưa, hãy quay lại Quan sát trạng thái.

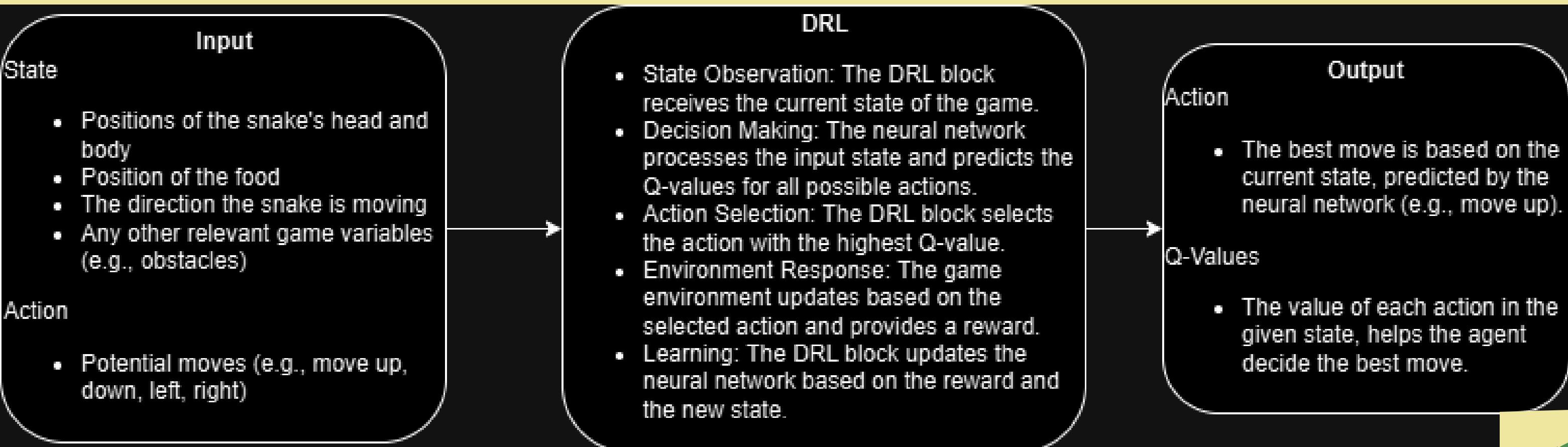
Bước 6: Điều chỉnh tham số

- Điều chỉnh các tham số và cải thiện mô hình nếu cần



DRL HOẠT ĐỘNG THẾ NÀO?

- Quan sát trạng thái: Khối DRL nhận trạng thái hiện tại của trò chơi
- Ra quyết định: Mạng nơ-ron xử lý trạng thái đầu vào và dự đoán giá trị Q cho tất cả các hành động có thể
- Lựa chọn hành động: Khối DRL chọn hành động có giá trị Q cao nhất
- Phản hồi môi trường: Môi trường trò chơi cập nhật dựa trên hành động đã chọn và cung cấp phần thưởng
- Học: Khối DRL cập nhật mạng nơ-ron dựa trên phần thưởng và trạng thái mới





A vibrant illustration of a medieval castle's courtyard. In the center is a large white speech bubble containing the text. To the left, a tiger stands near a green tree and a stone wall. To the right, a lion stands near a blue pond and a grey dragon. The background features grey stone pillars and small green bushes.

**THANK YOU
FOR LISTENING**