Snake game

Học tăng cường là một lĩnh vực của học máy liên quan đến cách các tác nhân phần mềm nên thực hiện các hành động trong một môi trường để tối đa hóa khái niệm phần thưởng tích lũy.

* Học tăng cường là dạy một tác nhân phần mềm cách cư xử trong một môi trường bằng cách cho nó biết nó đang làm tốt như thế nào.
* tác nhân: trình phát máy tính
* môi trường
* trao phần thưởng cho tác nhân
* cho nó biết nó đang làm tốt như thế nào và sau đó dựa trên phần thưởng của chúng, nó sẽ cố gắng tìm ra hành động tiếp theo tốt nhất.

Sử dụng học sâu (học q sâu): sử dụng mạng nơ-ron sâu để dự đoán các hành động

* tính toán trạng thái rồi dựa trên trạng thái đó chúng ta tính toán hành động tiếp theo.
* nhận được phần thưởng, trò chơi sẽ báo cáo điểm số.
* lưu trữ trạng thái mới và trạng thái cũ, trạng thái trò chơi và điểm số

Agent:

* Game
  + Use pygame

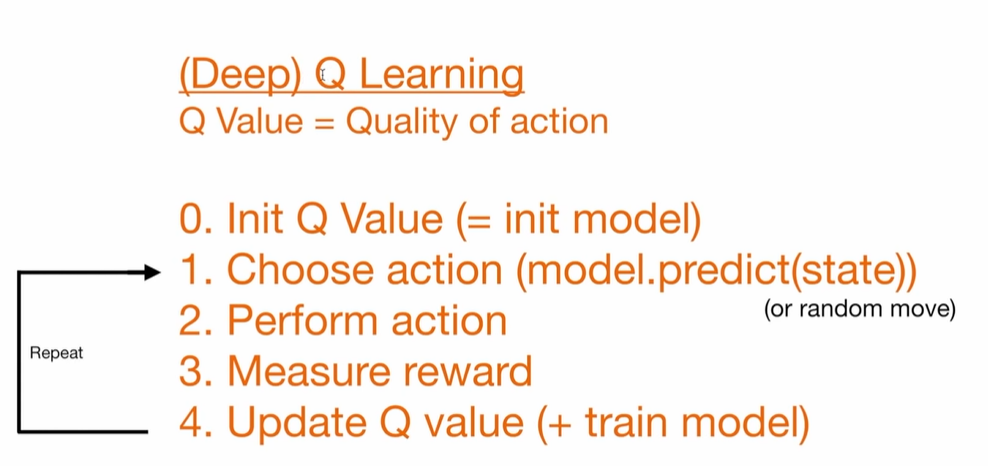
Play\_step(action)

* + - Reward, game\_over, score
* Model (PyTorch)
  + Linear\_QNet (DQN)

Model.predict(state)

* + - Action
* Training:
  + State = get\_state(game)
  + Action = get\_move(state):
    - Model.predict()
  + Reward, game\_over, score = game.play\_step(action)
  + New\_state = get\_state(game)
  + Remember
  + Model.train()
* Reward:
  + Eat food: +10
  + Game over: -10
  + Else: 0
* Action:
  + [1, 0, 0]: Straight
  + [0, 1, 0]: right turn
  + [0, 0, 1]: left turn
* State: tell the snake information about the game that it knows about the environment
  + Have 11 values:
    - Danger straight, danger right, danger left
    - Direction left, direction right, direction up, direction down
    - Food left, food right, food up, food down

(Deep) Q Learning



Bellman Equation

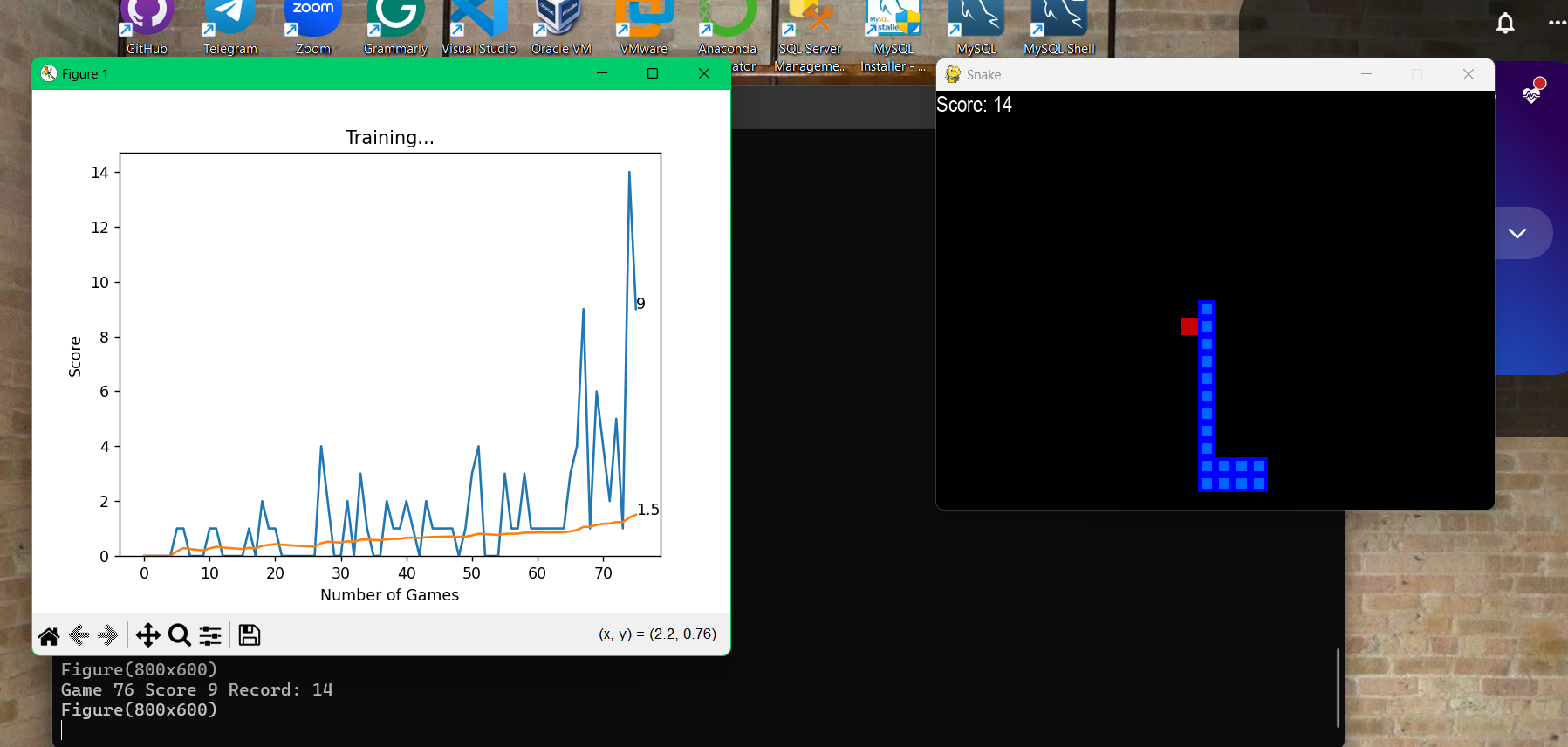
* : new Q value for that state and that action
* : current Q value
* : learning rate
* : reward for taking that action at that state
* : discount rate
* : max expected future reward given the new s’ and all possible actions at that new state

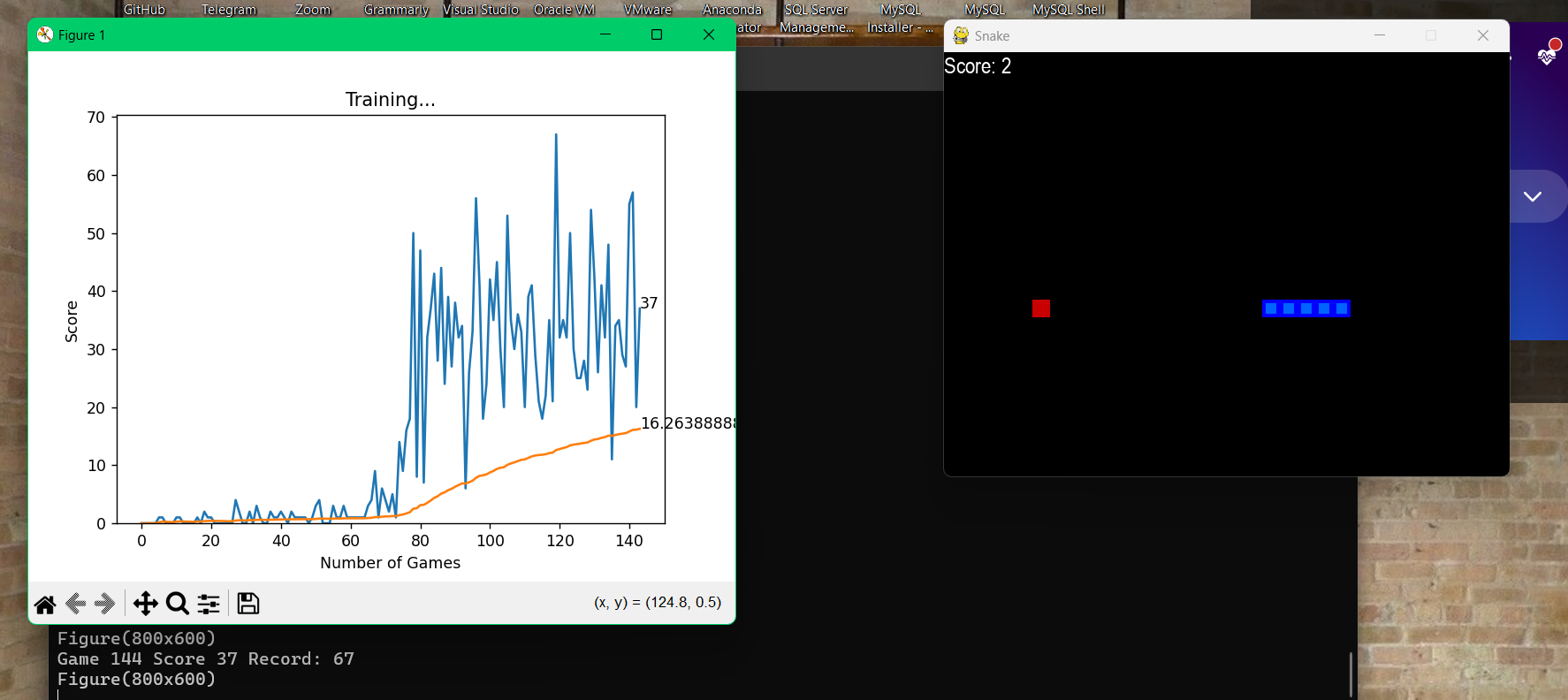
Q Update rule simplified

Loss function

Game

* Conda create -n snake\_game python=3.10
* Conda activate snake\_game





* Có 5 khối: đâu là epsilon, đâu là thuật toán tham lam, ảnh hưởng tới nhau ntn?