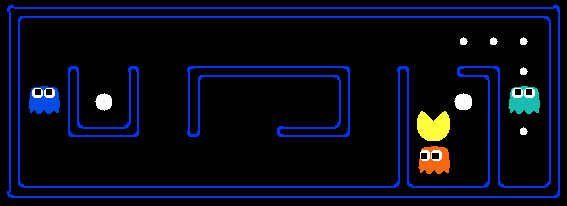
Reinforcement Learning and Pacman

1. **Giới thiệu**

Trong bài tập này, phần đầu tiên sẽ làm quen với thuật toán valueIteration và Qlearning trong môi trường đơn giản (GridWorld) (task1 và 2)

Sau đó, sẽ áp dụng Qlearning để học ra cách chơi của pacman trong môi trường thật bằng Reinforcement Learning (task 3 và 4).

Ví dụ cho hình: policy sẽ được học để quyết định pacman đi thẳng hay ăn power up?



Code-base được lấy từ <http://ai.berkeley.edu/reinforcement.html>

Phần code bạn cần update:

*valueIterationAgents.py*: Agent sử dụng thuật toán value iteration để giải quyết các bài toán MDP

*qlearningAgents.py*: Qlearning agent sử dụng Qlearning để giải bài toán GridWorld và Pacman

Các file có thể

1. **Task 1: Value Iteration trên GridWorld: 10 điểm**

Trong bài toán sử dụng value iteration, các bạn sử dụng thuật toán offline để update hàm value và policy đi kèm. Thuật toán cần tham số đầu vào là số lượng các bước lặp (sử dụng parameters -i).

Các bạn cần implement 2 thuật toán:

computeActionFromValues(state): Tìm ra hành động tốt nhất của 1 state

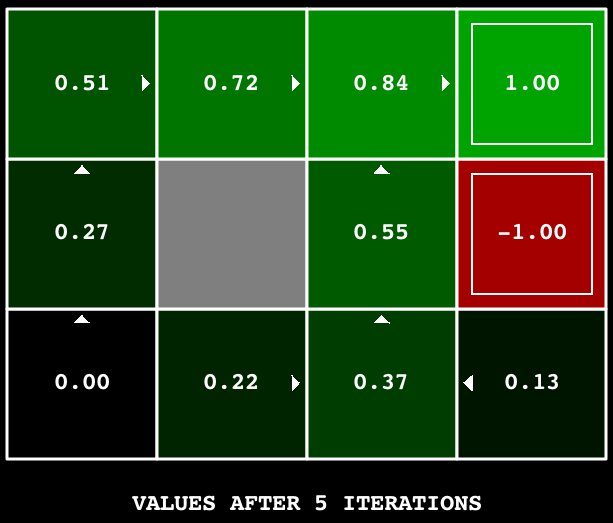
computeQValueFromValues(state, action): Trả về giá trị Qvalue của 1 cặp state action

Các bạn có thể test kết quả bằng dòng lệnh:

python autograder.py -q q1

Sử dụng GUI để test, với các tham số thích hợp:

python gridworld.py -a value -i 100 -k 10

Ví dụ, kết quả đúng của python gridworld.py -a value -i 5

Chấm điểm: chương trình sẽ được test bằng giá trị của V, và Q sau một số lần lặp và lúc giá trị của V hội tụ.

Tham khảo thêm q1 trên <http://ai.berkeley.edu/reinforcement.html>

1. **Task 2: Q-Learning trên GridWorld (10 điểm)**

Q-learning được dùng trên các môi trường mà mô hình MDP quá phức tạp và không thể biểu diễn được đầy đủ (do thiếu thông tin, agent buộc phải tự exploration và build mô hình MDP của mình).

Agent sử dụng Q-learning sẽ tự xây dựng hàm V và Q-values bằng cách thử. Trong task này, các hàm phải cài đặt ở trong *qlearningAgents.py*  gồm có computeValueFromQValues, getQValue, và computeActionFromQValues.

Lệnh để visualize:

python gridworld.py -a q -k 5 –m

Lệnh để chạy test:

python autograder.py -q q4

1. **Task 3: Q-Learing trên pacman (20 điểm)**

**part 1: 10 Điểm**

Sử dụng*qlearningAgents.py*  trong bài trên để huấn luyện trên môi trường đơn giản của pacman. Lệnh để test

python pacman.py -p PacmanQAgent -x 2000 -n 2010 -l smallGrid

Câu lệnh trên có nghĩa: huấn luyện pacman agent với 2000 iterations trên smallGrid

Chú ý: PacmanQAgent được định nghĩa với các giá trị phù hợp hơn cho bài toán pacman. Có thể test lại bằng cách sử dụng options –a:

-a epsilon=0.1,alpha=0.3,gamma=0.7

Điểm được chấm theo lệnh: python autograder.py -q q7

Chương trình sẽ train trong 2000 lượt chơi và test trên 100 lượt. Điểm sẽ cho theo số trận thắng trong 100 lượt chơi

**Part 2: 10 Điểm**

Update: Yêu cầu implement thuật toán xấp xỉ Qlearning, điểm đc cho theo score đạt đc khi training với 50 games (10\*your\_score/MAX\_SCORE)

Lệnh test: python autograder.py -q q8

1. **Nộp bài**

Chia theo nhóm từ 3 đến 5 thành viên.

Nộp bài: các bạn chỉ cần nộp 2 file valueIterationAgents.py và qlearningAgents.py qua website môn học.

Thời gian: trước 10/12.