Achieving mouse-level strategic evasion performance using real-time computational planning

German Espinosa, Gabrielle E. Wink, Alexander T. Lai,
Daniel A. Dombeck, Malcolm A. MacIver

Outline

Introduction

Method

MAP

Cognitive map

Predator

prey

Experiments

Result

Discusion

Introduction

Planning is the brain's ability to imagine possible futures and make them come true.

Animals with planning ability perform complex behaviors.

The purpose of this research is to develop an agent capable of strategic avoidance on the same level as living animals.

Method

Uses a spatial planning algorithm called TLPPO (Trajectories over Locations where Planning Pays Off)

TLPPO reduces the choice of destinations by pre-setting the locations where planning pays off, enabling real-time performance.

Use POMCP (Partially Observable Monte Calro Planning) for comparison.

POMCP is a method that develops a new trajectory each time the user moves, and determines the destination based on the statistics of the number of visits and the reward value that can be obtained on the spot.

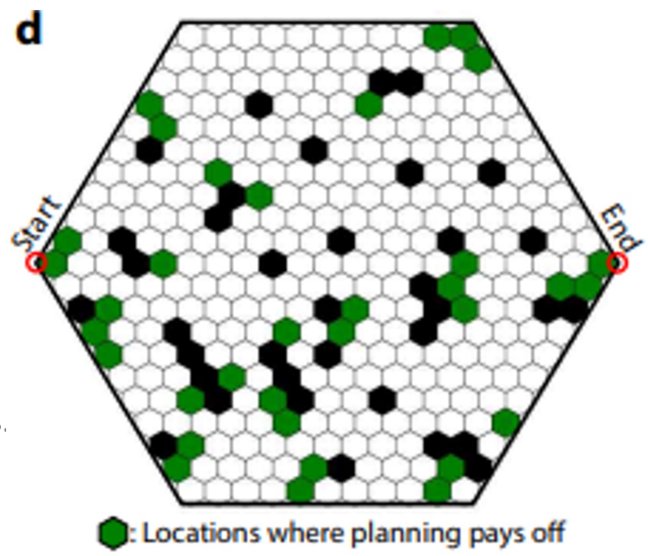
MAP

The map is made up of hexagons with obstacles placed inside.

The starting position of the prey is on the left, and the predator is placed on the goal side of the right.

The black part is an obstacle

The green areas are where agents can improve their planning efforts.



Cognitive map

Cognitive map is the brain's understanding of the surrounding space (MAP).

By forming a cognitive map, animals can capture and evade objects more smoothly and increase their success rate.

mouse were made to generate cognitive maps.

Predator

In the experiment, a robot is used as a predator, and if the robot gives an air puff to the prey, it is considered to have preyed.

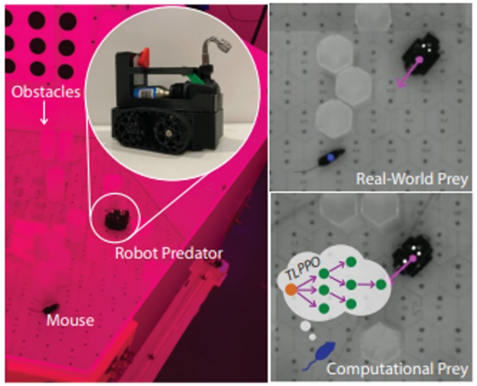
The algorithm is like this.

However, if the robot hits an obstacle and cannot return to its orbit, the joystick control intervenes.

Also, use white noise to cover the sound of the robot moving.

Algorithm 1 Autonomous robot predator behavior

```
while experiment is running do
    Find spawn cell S
    Move robot to S
    while episode is running do
        if mouse is visible then
            Move robot to last seen mouse cell
        else if mouse is not visible then
            Find cells not visible to robot
            Randomly select a non-visible cell N
            Move robot to N
        end if
    end while
end while
```



Prey

Two real mouses were used in the experiment.

TLPPO and POMCP prey robots were prepared as agents.

The prey was made to generate a "cognitive map."

Experiment

Each agent runs for 100 episodes and survival rates are recorded.

Compare whether the agent can perform as well as the mouse in the experiment.

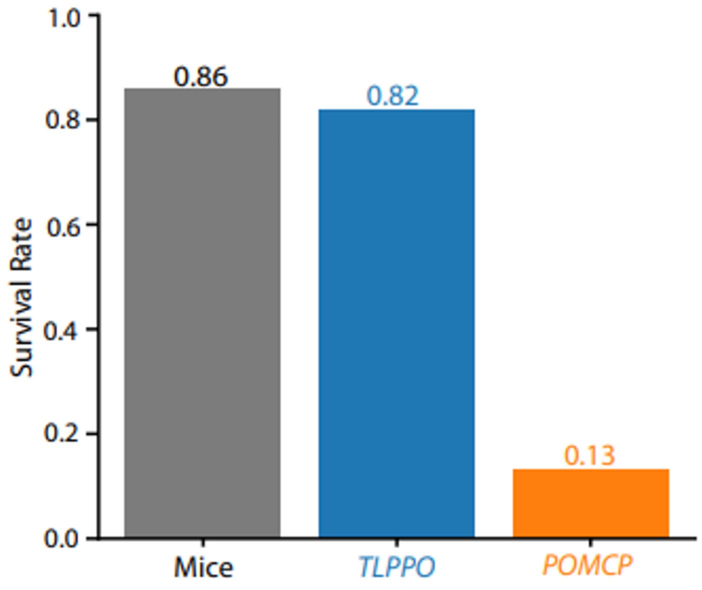
The actual survival rate of mouse is calculated as the average of two rmouses.

Result

The actual survival rate of the rats was 0.86.

The survival rate for POMCP was 0.13, and the survival rate for TLPPO was 0.82.

The survival rate of TLPPO was comparable to the performance of mouse.



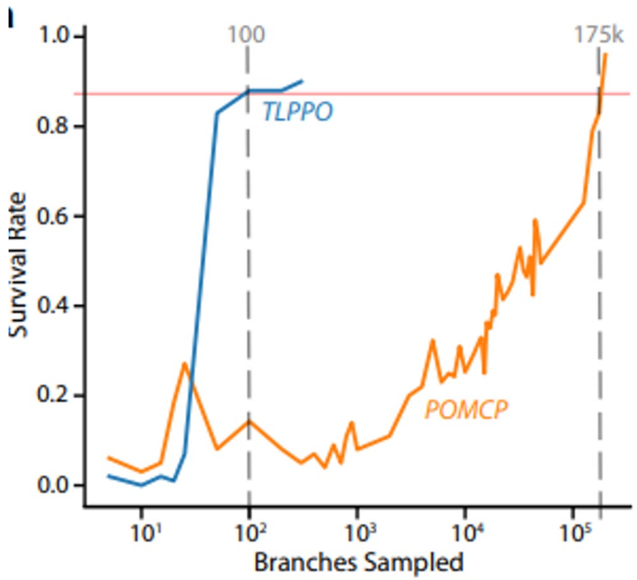
Result

This figure shows the growth graph of each agent.

This time, calculated the survival rate for 100 episodes.

POMCP was able to achieve the same survival rate after 10,000 episodes as TLPPO's 100 episodes.

Therefore, TLPPO was able to survive very efficiently.



Discussion

In this study, compared POMCP and TLPPO algorithms and showed their efficiency.

In real nature, prey cannot learn because they would die if they were caught by a predator. Therefore, it is difficult to apply to naturalistic problems.

Finally, able to demonstrate that it is possible to more efficiently increase survival rates in the extremely unstable environment of TLPPO by combining complex real-time interaction tasks and cognitive maps.

Thank you for attention