UIR: Game Theory in Robotics - Lab 1

David Milec

Czech Technical University in Prague

milecdav@fel.cvut.cz

December 6, 2021

Overview

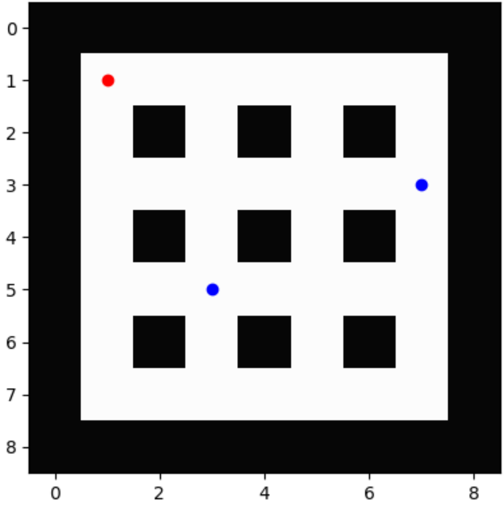
1. Game Theory
2. Pursuit Evasion Game
3. Two-player Zero-sum Game
4. Double Oracle
5. Stochastic Game
6. Value Iteration in Stochastic Games

Game Theory

- Framework studying strategies of players when the outcome of the actions depends on the actions of the other players.
- In the assignments we will focus on two-player zero-sum games.
- Optimal strategy in such games is described by Nash equilibrium.

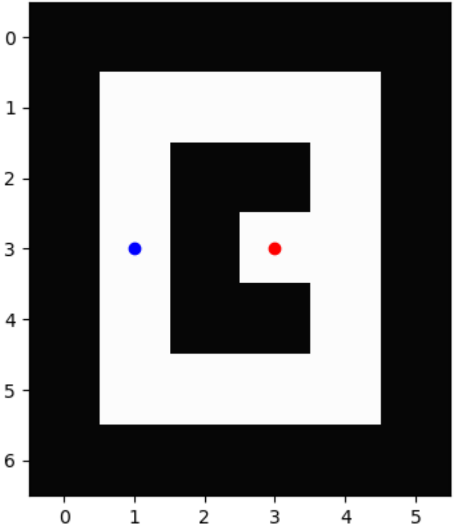
Pursuit Evasion Game

- Game is played in a grid environment.
- Players use simultaneous discrete moves.
- Both players have perfect information about the environments and the other player.
- Evader (red) gets payoff for escaping for a fixed amount of steps.
- Pursuers (blue) get payoff for catching the evader.



Heuristic approaches

- Doing a move in such a way that I end in a space closest/furthest to/from the opponent.
- Euclidean distance does not work for pursuer even against a stationary opponent.
- Closest path is better but does not work with more pursuers in a circular environment.



Greedy Policy

- First task: t4a-greedy
- Implement player that will use greedy strategy - pursuers will move towards closest evader and evader will go to a place that is as far as possible from the closest pursuer.
- https://cw.fel.cvut.cz/wiki/courses/b4m36uir/hw/t4a-greedy

Two-player Zero-sum Game

- Strategy sets M and N
- Utility matrix $\mathbf{U} = [c_{ij}]_{i \in M, j \in N}$
- Represented as matrix and also called Matrix game or Normal-form Game (NFG)
- Example: $|M| = 2, \quad |N| = 3, \quad \mathbf{U} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 0 \end{bmatrix}$
- Row player with 2 actions is maximizing and column player with 3 actions is minimizing

Double Oracle

- Used to find Nash Equilibrium of zero-sum two-player game (in our example normal form game).
- Pick randomly one action for each players and form restricted subgame using those actions.
- In each iteration solve the subgame and find the best responses to the current strategy profile in the subgame for both players and add them to the restricted subgame, then solve the subgame again.
- We stop when we encounter iteration where both best responses are already in the subgame.

Try Double Oracle algorithm on the following matrix game

	A	B	C	D	E
V	-8	9	0	7	-6
W	6	9	6	5	6
X	1	-8	3	8	7
Y	5	2	6	-5	2
Z	4	3	3	0	8

Double Oracle

BRs $= \{W, A\}$ - A already in the subgame so we add W
Solution of the subgame gives A and W with probability 1

	A	B	C	D	E
V	-8	9	0	7	-6
W	6	9	6	5	6
X	1	-8	3	8	7
Y	5	2	6	-5	2
Z	4	3	3	0	8

Double Oracle

BRs $= \{W, D\}$ so we add D to the subgame and solve it

	A	B	C	D	E
V	-8	9	0	7	-6
W	6	9	6	5	6
X	1	-8	3	8	7
Y	5	2	6	-5	2
Z	4	3	3	0	8

Double Oracle

Strategy for Player 1: $X(V) = \frac{1}{16}, X(W) = \frac{15}{16}$

Strategy for Player 2: $Y(A) = \frac{1}{8}, Y(D) = \frac{7}{8}$

	A	B	C	D	E	
V	-8	9	0	7	-6	5.125
W	6	9	6	5	6	5.125
X	1	-8	3	8	7	7.125
Y	5	2	6	-5	2	-3.75
Z	4	3	3	0	8	0.5
	$5\frac{1}{8}$	9	$5\frac{5}{8}$	$5\frac{1}{8}$	$5\frac{2}{8}$	

Double Oracle

BRs $= \{X, (A, D)\}$ so we add X to the subgame and we solve it

	A	B	C	D	E	
V	-8	9	0	7	-6	5.125
W	6	9	6	5	6	5.125
X	1	-8	3	8	7	7.125
Y	5	2	6	-5	2	-3.75
Z	4	3	3	0	8	0.5
	$5\frac{1}{8}$	9	$5\frac{5}{8}$	$5\frac{1}{8}$	$5\frac{2}{8}$	

Double Oracle

Strategy for Player 2: $S_1(A) = \frac{3}{8}, S_1(D) = \frac{5}{8}$

Strategy for Player 1: $S_2(V) = 0, S_2(W) = \frac{7}{8}, S_2(X) = \frac{1}{8}$

	A	B	C	D	E	
V	-8	9	0	7	-6	1.375
W	6	9	6	5	6	5.375
X	1	-8	3	8	7	5.375
Y	5	2	6	-5	2	-1.25
Z	4	3	3	0	8	1.5
	$5\frac{3}{8}$	$6\frac{7}{8}$	$5\frac{5}{8}$	$5\frac{3}{8}$	$6\frac{1}{8}$	

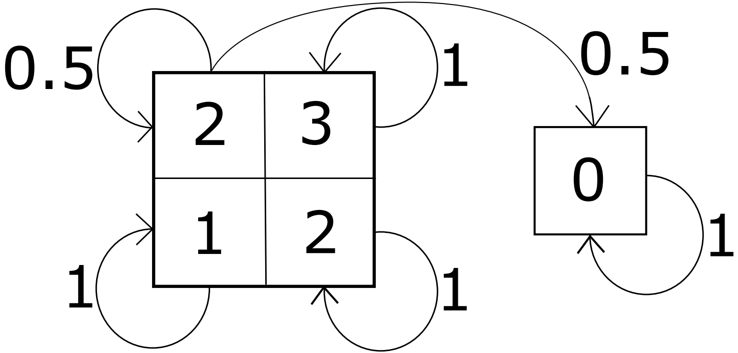
Double Oracle

All best responses are already in the subgame
so we stop and we have a Nash equilibrium

	A	B	C	D	E	
V	-8	9	0	7	-6	1.375
W	6	9	6	5	6	5.375
X	1	-8	3	8	7	5.375
Y	5	2	6	-5	2	-1.25
Z	4	3	3	0	8	1.5
	$5\frac{3}{8}$	$6\frac{7}{8}$	$5\frac{5}{8}$	$5\frac{3}{8}$	$6\frac{1}{8}$	

Stochastic Game

- Strategy sets M and N
- The set S of states
- A transition function: $T : S \times M \times N \rightarrow \Delta_S$
- A rewards function: $R : S \times M \times N \rightarrow \mathbb{R}$



Value Iteration

- Value iteration in stochastic games in an adaptation of value iteration used to solve MDPs.
- It stores all the values for all possible states of the game.
- Value iteration iteratively updates those values based on possible actions in each state, solving matrix game created from next state values.
- Finally, the value iteration uses computed values to compute the strategy.

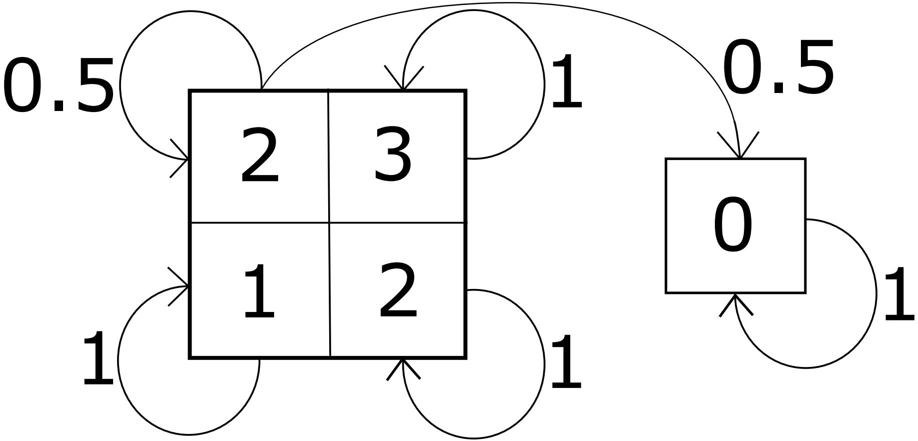
Value Iteration

S is the state space, $v : S \to \mathbb{R}$ is value in each state, $\mathcal{A}$ is set of all combinations of actions and $A : S \to \mathcal{A}$ is a function returning all possible action tuples available in a given state. Q is a matrix game created for each state in each iteration, $r : S \times \mathcal{A} \to \mathbb{R}$ is immediate payoff and $T : S \times \mathcal{A} \to S$ is a transition function. γ is discounting constant.

$$
\begin{aligned}
&\forall s \in S \quad \text{initialize} \quad v(s) = 0 \quad \text{and until v converges} \\
&\forall s \in S \\
&\qquad \forall (a_1, a_2) \in A(s) \\
&\qquad\qquad Q(a_1, a_2) = r(s, a_1, a_2) + \gamma \sum_{s' \in S} T(s, a_1, a_2) v(s') \\
&\qquad v(s) = \max_x \min_y xQy
\end{aligned}
$$

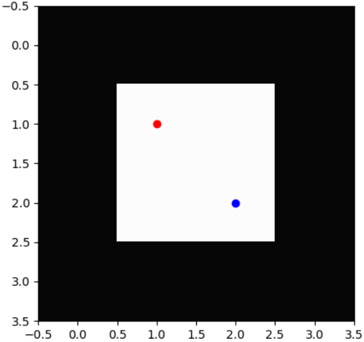
Value Iteration

Try Value iteration on the example game with $\gamma = 0.5, \gamma = 0.9$ and $\gamma = 1$.



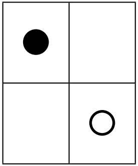
Pursuit Evasion Game as Stochastic Game

- On the example game with one pursuer and one evader we will assume that move is successful with probability 75%, otherwise the agent does not move. Create a stochastic game representation of it. Hint: Put all positions equal under rotation and mirroring in one state.

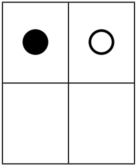


Pursuit Evasion Game as Stochastic Game

We have states 1 and 2 which contains all possible configurations if we allow rotation and mirroring. Then we have state 3 which is absorbing state for the catch. Transition probabilities are in the tables below.



State 1



State 2

	D	R
L	$\left(\frac{1}{16}, \frac{6}{16}, \frac{9}{16}\right)$	$\left(\frac{10}{16}, \frac{6}{16}, 0\right)$
U	$\left(\frac{10}{16}, \frac{6}{16}, 0\right)$	$\left(\frac{1}{16}, \frac{6}{16}, \frac{9}{16}\right)$

	D	R
L	$\left(\frac{3}{16}, \frac{10}{16}, \frac{3}{16}\right)$	$\left(0, \frac{1}{16}, \frac{15}{16}\right)$
D	$\left(\frac{6}{16}, \frac{10}{16}, 0\right)$	$\left(\frac{3}{16}, \frac{10}{16}, \frac{3}{16}\right)$

Value Iteration in Homework

- Your homework is to implement value iteration for the pursuer evader game shown at the beginning.
- In the homework we either catch, thus ending the game, and we get 1 or we move to single new state, resulting in:

$$
\begin{aligned}
&\forall s \in S \quad \text{initialize} \quad v(s) = 0 \quad \text{and until v converges} \\
&\forall s \in S \\
&\qquad \forall (a_1, a_2) \in A(s) \\
&\qquad\qquad Q(a_1, a_2) = \text{if catch 1 else } (\gamma v(s')) \\
&\qquad v(s) = \max_x \min_y xQy
\end{aligned}
$$

- $\max_x \min_y xQy$ requires you to formulate and solve a linear program to find a Nash equilibrium.

Linear Program to Find Nash Equilibrium

Create Linear programs for both players in this game

$$\mathbf{U} = \begin{bmatrix} 1 & 2 & 3 \\ 6 & 5 & -2 \end{bmatrix}$$

Primal

$$\begin{aligned} \text{Maximize} \quad & v \\ \text{subject to} \quad & \mathbf{x}^\mathsf{T}\mathbf{U}\mathbf{e}_j \geq v, \forall j \in N \\ & x_i \geq 0, \forall i \in M \\ & \sum_{i \in M} x_i = 1 \\ & v \in \mathbb{R} \end{aligned}$$

Dual

$$\begin{aligned} \text{Minimize} \quad & v \\ \text{subject to} \quad & \mathbf{y}^\mathsf{T}\mathbf{U}^\mathsf{T}\mathbf{e}_j \leq v, \forall j \in M \\ & y_i \geq 0, \forall i \in N \\ & \sum_{i \in N} y_i = 1 \\ & v \in \mathbb{R} \end{aligned}$$

Linear Program to Find Nash Equilibrium

$$\mathbf{U} = \begin{bmatrix} 1 & 2 & 3 \\ 6 & 5 & -2 \end{bmatrix}, \quad \bar{x} = \left(\frac{4}{5}, \frac{1}{5}\right), \bar{y} = \left(\frac{1}{2}, 0, \frac{1}{2}\right), \bar{v} = 2$$

Primal

$$\begin{aligned} \text{maximize} \quad & v \\ \text{s.t.} \quad x_1 + 6x_2 &\geq v \\ 2x_1 + 5x_2 &\geq v \\ 3x_1 - 2x_2 &\geq v \\ x_1 + \ \ x_2 &= 1 \\ x_1, x_2 &\geq 0 \\ v &\in \mathbb{R} \end{aligned}$$

Dual

$$\begin{aligned} \text{minimize} \quad & v \\ \text{s.t.} \quad y_1 + 2y_2 + 3y_3 &\leq v \\ 6y_1 + 5y_2 - 2y_3 &\leq v \\ y_1 + \ \ y_2 + \ \ y_3 &= 1 \\ y_1, y_2, y_3 &\geq 0 \\ v &\in \mathbb{R} \end{aligned}$$

The End