Game theory - lab 2

David Milec

Czech Technical University in Prague

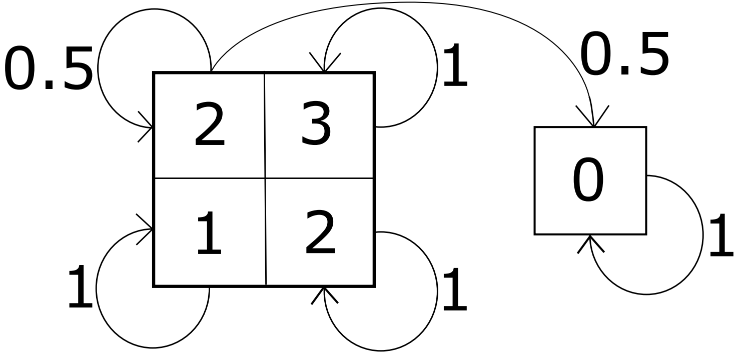
milecdav@fel.cvut.cz

December 12, 2022

Overview

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.

$$\forall s \in S \quad \text{initialize} \quad v(s) = 0 \quad \text{and until v converges}$$

$$\forall s \in S$$

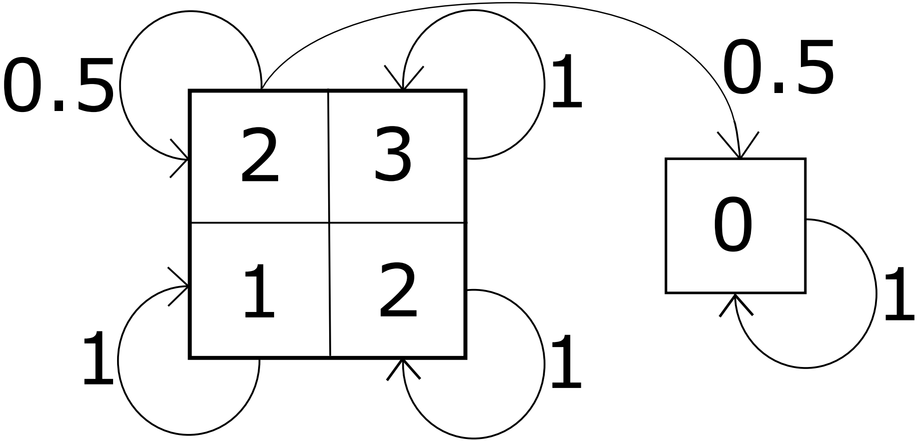
$$\quad \forall (a_1, a_2) \in A(s)$$

$$\qquad Q(a_1, a_2) = r(s, a_1, a_2) + \gamma \sum_{s' \in S} T(s, a_1, a_2) v(s')$$

$$\quad v(s) = \max_x \min_y xQy$$

Value Iteration

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



Value iteration solution $\gamma = 0.5$
(u_{ij} are values of the matrix game we need to solve)

iteration	$V(s_1)$	$V(S_2)$	u_{11}	u_{21}	u_{12}	u_{22}
1	2	0	2	1	3	2
2	2.5	0	2.5	2	4	3
3	2.625	0	2.625	2.25	4.25	3.25
4	2.656	0	2.656	2.313	4.313	3.313
5	2.664	0	2.664	2.328	4.328	3.328
6	2.664	0	2.664	2.328	4.328	3.328
7	2.666	0	2.666	2.333	4.333	3.333

Value iteration solution $\gamma = 0.9$

iteration	$V(s_1)$	$V(S_2)$	u_{11}	u_{21}	u_{12}	u_{22}
1	2	0	2	1	3	2
2	2.9	0	2.9	2.8	4.8	3.8
3	3.61	0	3.305	3.61	5.61	4.61
4	4.249	0	3.625	4.249	6.249	5.249
5	4.824	0	3.912	4.824	6.824	5.824
6	5.342	0	4.171	5.342	7.342	6.342
7	5.808	0	4.404	5.808	7.808	6.808
...	...	0	...	...	...	...
137	10	0	6.5	10	12	11

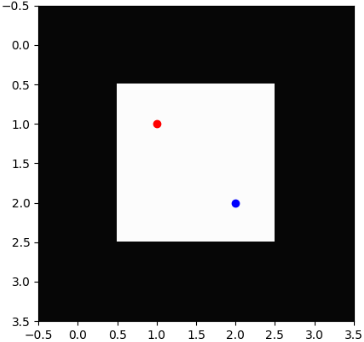
Value iteration solution $\gamma = 1$

iteration	$V(s_1)$	$V(S_2)$	u_{11}	u_{21}	u_{12}	u_{22}
1	2	0	2	1	3	2
2	3	0	3	3	5	4
3	4	0	3.5	4	6	5
4	5	0	4	5	7	6
5	6	0	4.5	6	8	7
6	7	0	5	7	9	8
7	8	0	5.5	8	10	9
...	...	0	...	...	...	...
100	102	0	52.5	102	104	103

Will go up to infinity.

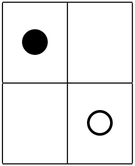
Pursuit Evasion Game as Stochastic Game

- 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.

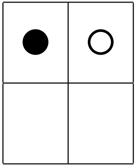


Pursuit Evasion Game as Stochastic Game

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:

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

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