Kriangkrai Chaonithi
57070701702

AI-Project: MDP Simulation on Web Browser

Markov Decision Process (MDP) Simulator on Web Browser is a web application written on JavaScript as a core and Angular Material for user interface. The application can run on most modern web browsers on computers, tablets, also smart phones. It is open source and available at https://www.spicydog.org/project/mdp-simulator/.

Users can adjust several configuration as shown in figure 1 including field, rate of convergence (γ), and uncertainty. The size of field can adjust in width and height from 2 to 36. The change of size immediately displayed on the field configuration that user can change type of each position in the field by clicking to the box to toggle. The colors represent the type: light-gray for blank, green for goal, red for trap, and black for obstacle. Next is the γ configuration that will be used in the equation 1. Last, users can also configure uncertainty of moves to the program. After finish configuration, the button ready will setup the MDP environment and ready for a simulation.

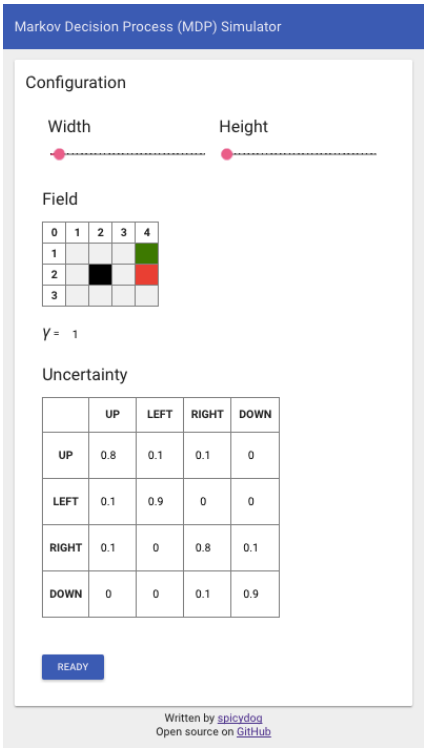

Figure 1 Configuration Interface

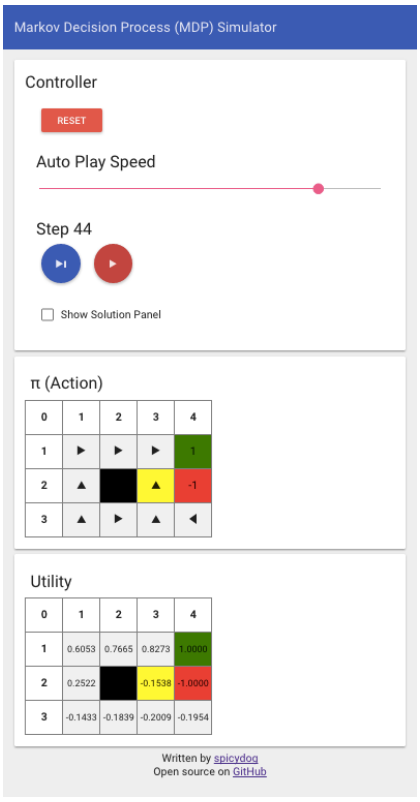

Figure 2 Simulation Interface

The simulation interface is shown in the figure 2. The controller panel is one top of the screen that includes reset button for users to click and go back to configuration panel. The auto play speed adjusts the interval of each step during the simulation. The circle buttons control the simulation process. The blue button run the simulator step by step and the red button runs the simulator automatically until the users press it again to pause. Lastly, the show solution checkbox shows the panel that displays the calculations on each step.

The simulation interface also includes policy (π) and utility table. The size, type, and color of fields are as same as the configuration excepts the yellow box that represents the current position of the agent. The policy table represents the move that agent moved when it was at that position. The utility table represents the score of each position in the field. When the simulator runs, users can see changes to these interface that shows the results of MDP algorithm.

In the calculation part of simulator, we follow the MDP algorithm with equation 1 for utility calculation and equation 2 for action that the agent will take. The rewards of each position is not allowed to adjust on this version and it is set as -0.04 of the normal block. The agent will be randomly move to the new location when it reaches the terminal block. We also did not apply terminate condition for simulator but the users can see the behavior of the simulator when it each the equilibrium state.

$$U(s) = R(s) + \gamma \max_{a} \sum_{s'} T(s,a,s')U(s')$$

Equation 1 Utility

$$\pi(s) = \operatorname{argmax}_{a} \sum_{s'} T(s,a,s')U(s')$$

Equation 2 Action