

This handout includes space for every question that requires a written response. Please feel free to use it to handwrite your solutions (legibly, please). If you choose to typeset your solutions, the `README.md` for this assignment includes instructions to regenerate this handout with your typeset L<sup>A</sup>T<sub>E</sub>X solutions.

---

1.a

$$V_{minimax}(s, d) = \begin{cases} Utility(s) & \text{if } isEnd(s) \\ Eval(s) & \text{if } d = 0 \\ \max_{a \in Actions(s)} V_{minimax}(Succ(s, a), d) & \text{if } Player(s) = \text{Pacman} \\ \min_{a \in Actions(s)} V_{minimax}(Succ(s, a), d - \frac{1}{k}) & \text{if } Player(s) = \text{Ghost} \end{cases}$$

3.a

$$\Pi(s, a) = \frac{1}{|Actions(s)|}$$

$$V_{expectimax}(s, d) = \begin{cases} Utility(s) & \text{if } isEnd(s) \\ Eval(s) & \text{if } d = 0 \\ \max_{a \in Actions(s)} V_{expectimax}(Succ(s, a), d) & \text{if } Player(s) = \text{Pacman} \\ \sum_{a \in Actions(s)} \Pi(s, a) \cdot V_{expectimax}(Succ(s, a), d - \frac{1}{k}) & \text{if } Player(s) = \text{Ghost} \end{cases}$$

## 5.a

The MinimaxAgent's evaluation function returns the same value for all actions, so it chooses the first one. This action happens to be toward the nearest ghost. This is because Minimax assumes the opponent is playing optimally, which results in states that have no utility for the player. Expectimax changes this assumption, which results in actions with potential states that the player has utility, prioritizing those actions.

## 5.b

We could make MinimaxAgent behave more like Expectimax by changing the evaluation function to consider pellet and capsule distance. This would work because actions that move toward the food would provide utility, instead of all actions for Pacman having no utility due to Minimax selecting the optimal action for Ghosts.

### 5.c

An example of reward hacking would be self-driving cars with an objective function for shortest trip time. This objective function could have side effects such as law breaking, damage to the vehicle, or injury to people. It would choose to go over/through obstacles at the fastest possible rate, and would over-exert the vehicle in pursuit of reaching the destination as fast as possible. A better reward hack for this objective function would be exploding the car to make the passenger reach their destination the fastest.