


# CMPSCI 687 Pop Quiz 1

**Instructions:** You have 10 minutes to complete this quiz. This quiz is **closed** notes—do not use your notes or a laptop. Do not discuss problems with your neighbors until after everyone has handed in their quiz.

Recall the gridworld described in class:

Start $s = 1$	$s = 2$	$s = 3$	$s = 4$	$s = 5$
$s = 6$	$s = 7$	$s = 8$	$s = 9$	$s = 10$
$s = 11$	$s = 12$		$s = 13$	$s = 14$
$s = 15$	$s = 16$		$s = 17$	$s = 18$
$s = 19$	$s = 20$	$R = -10$ $s = 21$ 	$s = 22$	$R = +10$ $s = 23$ Goal

**Actions:**

attempt\_up  
attempt\_down  
attempt\_left  
attempt\_right

When the agent attempts to move in a direction:

The agent succeeds,  $p = 0.8$   
The agent veers  $90^\circ$  right,  $p = 0.05$   
The agent veers  $90^\circ$  left,  $p = 0.05$   
The agent stays in place,  $p = 0.1$

If the agent would ever hit a wall, it stays in its current position.

All unspecified rewards are zero.  
All specified rewards are for entering the state

Note: If the agent is in a state,  $s$ , takes an action,  $a$ , and transitions back to state  $s$ , we considering that “entering” the state  $s$ .

1. What is  $P(12, \text{attempt\_up}, 9)$ ?
2. What is  $R(21, \text{attempt\_right}, 21)$ ?
3. What is  $R(21, \text{attempt\_right}, 22)$ ?
4. If we change the reward function so that hitting a wall results in a reward of  $-1$ , then what would  $R(17, \text{attempt\_left}, 17)$  be?
5. If  $\gamma < 1$  and the gridworld is modified so that the agent’s actions always succeed, what is  $J(\pi^*)$  in terms of  $\gamma$ ?
6. What reward does the agent receive when it enters the absorbing state?