Quiz - 1

## Markov Decision Process

1. For a **markov process** with the following transition matrix,

|  | s1 | s2 | s3 |
| --- | --- | --- | --- |
| s1 | 0.1 | 0.2 | 0.3 |
| s2 | 0.7 | 0.5 | 0.4 |
| s3 | 0.6 | 0.8 | 0.9 |

What are the values of the following transition probabilities: **P(St+1=s1 | St=s3),** **P(St+1=s2 | St=s3) and P(St+1=s2 | St=s1, St-1=s1)**?

1. In an MDP, on which of the following does the state **St+1** depend on? (Choose one or more that apply)
   1. **St**
   2. **St-1**
   3. **At**
   4. **At-1**
   5. **Rt**
   6. **Rt+1**
2. What is the difference between “one step reward" and “long term return”?
3. What happens when the discount factor (𝛄) is set to 0 and when it is set to 1?
4. For episodic tasks, should the discount factor always be set to 0? Why or why not, give an example.
5. Is long term return **Gt** a random variable? Explain.