**Assignment 3**

**Nitzan Levy 322572488**

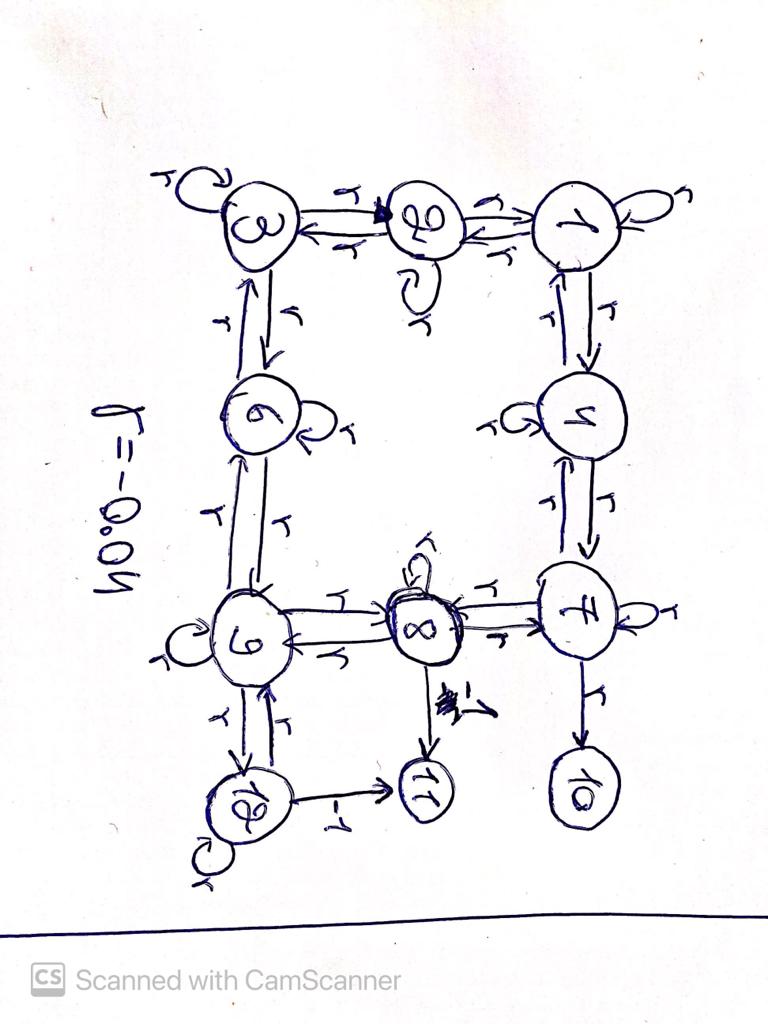
Question 1

* The Markov property meant that evolution of the Markov process in the future depends only on the present state and doesn’t on history.  
  A stochastic process is a process whose development depends on accidental factors.  
  A stochastic process has the Markov property if the conditional probability distribution of the future states of the process depends only upon the present state.
* A Markov Chain is a memory-less random process, a sequence of random states S1,S2… with the Markov property.  
  Markov chain is a tuple <S,P> :   
  S is a finite set of states.  
  P is a state transition probability matrix
* A Markov Decision Process (MDP) is a Markov reward process with decisions.   
  A MDP is a tuple <S, A, P, R, y>:  
  S is a finite set of states, A is a finite set of actions, P is a state transition, probability matrix, R is a reward function, y is a discount factor.
* Value iteration computes the optimal state value function by iteratively improving the estimate of V(s).
* Policy is agent’s behavior function.

Policy Iteration is an algorithm which help in learning the optimal policy which maximizes the long-term discounted reward.

Question 2

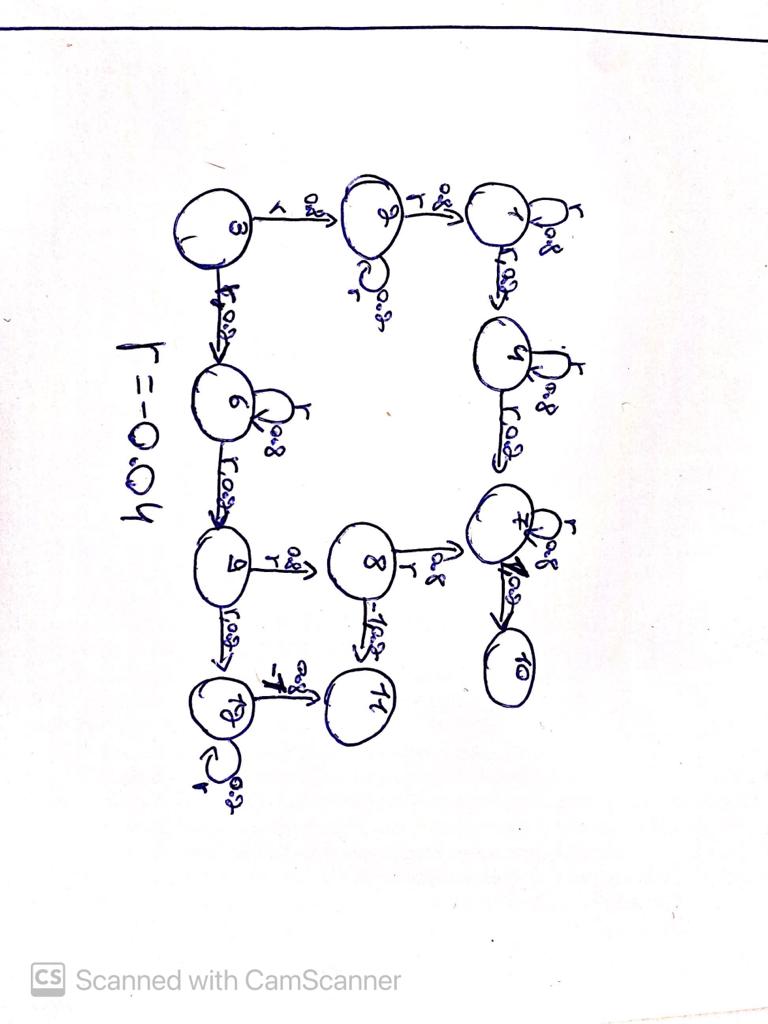
a.



1

An intuitive solution for the MDP is to take the path with the biggest weight. Weight of path calculated by the edges sum.   
For example: 3->2->1->4->7->10

b.



c.