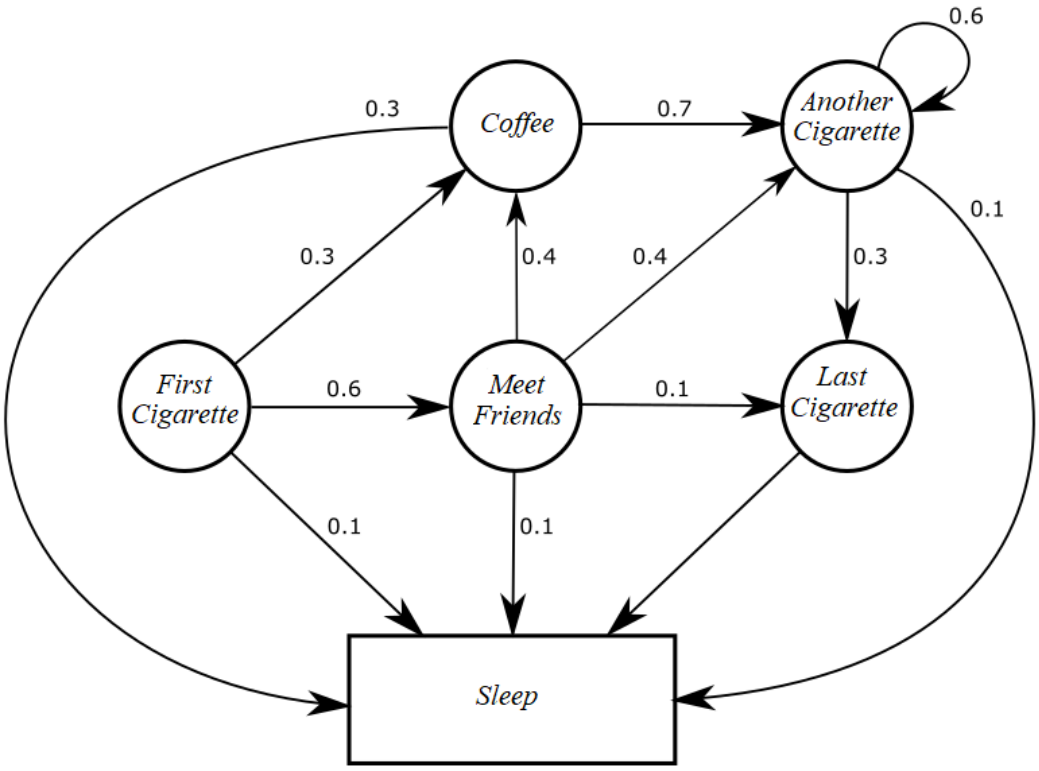
IE-456/556 & EEE-448/548
Reinforcement Learning and Dynamic Programming
Homework Assignment 1 – Due June 20 23:59

1. (40 points) The graph below represents a model of Sarah's afternoons when she returns home from work until she sleeps.



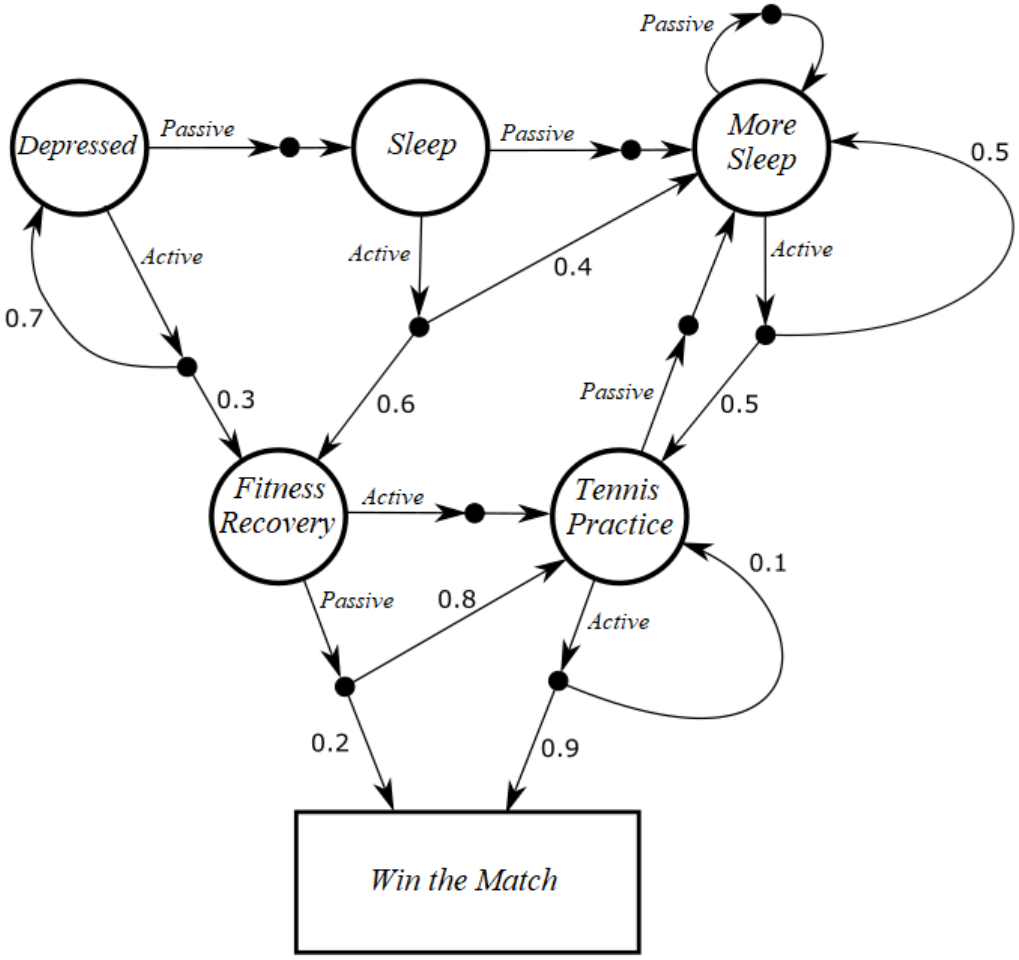
Consider the state space as {First Cigarette, Meet Friends, Coffee, Another Cigarette, Last Cigarette, Sleep} and the corresponding reward as {+1,+1,+2,+1,-3,0}.

(a) Construct the transition probability of the above model.

(b) Calculate the stationary probability distribution of the above model.

(c) Find the state-value function of each state using Bellman equation for $\gamma = 0.1$. Which state is the most promising in this case? Why do you think this the case? What state do you expect to be the most promising when $\gamma = 0.9$?

2. (60 points) Sarah is a professional tennis player and below you can find an MDP modeling one chapter of her professional career. In every state the choice is between two actions (being "Active" or "Passive").



Consider the state space as {Depressed, Sleep, More Sleep, Fitness Recovery, Tennis Practice, Win the Match} and the corresponding rewards as {-1,-1,-1,-1,-1,0} for both actions.

(a) Define the state transition probability matrices for both policies of "Active" and "Passive".

(b) Evaluate both the "Active" and "Passive" policies (find their state-value function of each state) for $\gamma = 0.9$.

(c) Given a randomized policy which has equal probability of taking the two actions in each state, calculate action-value function for the each state-action pair.

(d) Given a randomized policy which has α probability of taking the "Active" action in each state, plot the state-value function of the states "Fitness Recovery" and "Tennis Practice" for different values of α. What do we observe?