CSCD84: Artificial Intelligence:
Problem Set 3: MDP

UTORid:

First and Last Name:

By turning in this assignment, I agree by the University of Toronto honor code and declare that all of this is my own work.

An MDP with a single goal state ($S3$) is given below. $a1$, $a2$, and $a3$ represent the possible actions, and the pair of values next to each edge represent the probability and the cost of taking some action from a state and ending up at another one. Consider the state $S1$ for instance. Actions $a1$ and $a2$ are the two possible actions at this state and $0.75/2$ represent that with probability 0.75, action $a1$ will lead to state $S1$ and the cost of that is 2. [**NOTE:** In this example we are considering costs instead of rewards, hence it is a cost minimization problem.]

1. Given the expected cost for each state as $C(S1) = 7, C(S2) = 4.2$, and $C(S3) = 0$, calculate the optimal policy for state $S1$.
2. Suppose that we want to follow a policy where we pick action $a2$ in state $S1$ and action $a3$ in state $S2$. Calculate the expected cost of $S1$ and $S2$ for this policy.