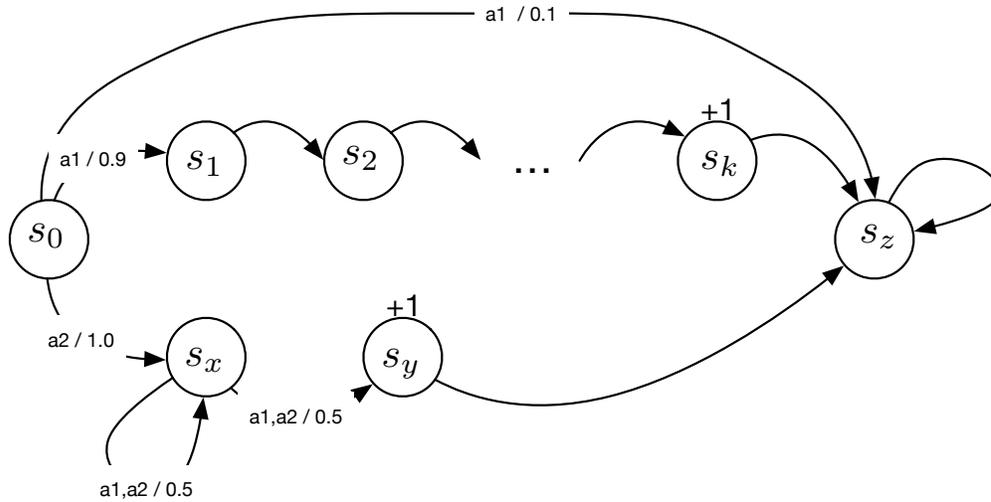


## Making Decisions Practical

5. (13 points) Consider the following MDP with  $k + 4$  states. There are two actions,  $a_1$  and  $a_2$ . Arrows with no labels represent a transition for both actions with probability 1. Arrows labeled  $a/p$  make the transition on action  $a$  with probability  $p$ . States with no label have reward 0. Two states have reward  $+1$ , obtained when taking an action in that state. There are  $k - 2$  states between  $s_1$  and  $s_k$ , with a deterministic transition on any action (so that once you are in  $s_1$  you are guaranteed to end up in  $s_k$  in  $k - 1$  steps).

We are interested in the infinite-horizon discounted values of some states in this MDP.



- (a) What is  $V(s_1)$  as a function of  $k$  when  $\gamma = 0$ ? \_\_\_\_\_
- (b) What is  $V(s_1)$  as a function of  $k$  when  $\gamma = 1$ ? \_\_\_\_\_
- (c) What is  $V(s_1)$  as a function of  $k$  when  $0 < \gamma < 1$ ? \_\_\_\_\_
- (d) What is  $V(s_x)$  when  $\gamma = 0$ ? \_\_\_\_\_
- (e) What is  $V(s_x)$  when  $\gamma = 1$ ? \_\_\_\_\_
- (f) What is  $V(s_x)$  when  $0 < \gamma < 1$ ? \_\_\_\_\_

Name: \_\_\_\_\_

- (g) Under what conditions on  $k$  and  $\gamma$  would we prefer to take action  $a_1$  in state  $s_0$ ? Write down a specific mathematical relationship.

