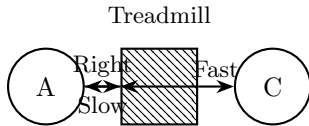


(a) Varying reward function (highlighted state has reward 10).



(b) Varying transition function: Treadmill pushes agent backward if slow.

Figure 1: Example MDPs with varying reward (left) and transition (right) functions.