



figureA diagram depicting example

MDPs where the reward or transition function varies. This diagram is used to illustrate how policies may vary with respect to changes in their reward (left) or transition functions (right). In the figure on the right, the box with arrows in it indicates a treadmill. If the agent is faster than the treadmill, it can pass over the treadmill. Otherwise, the treadmill would push the agent backwards.