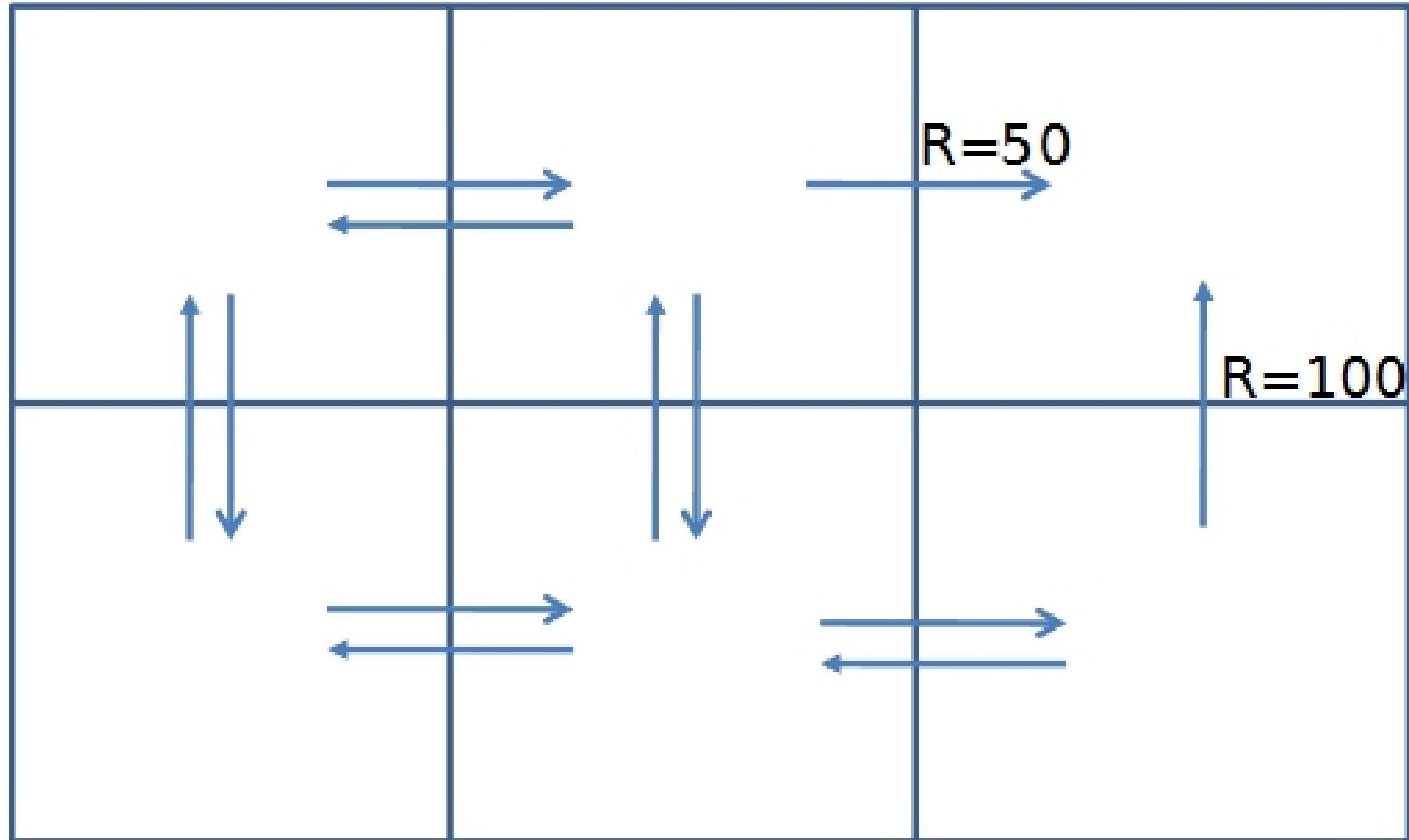


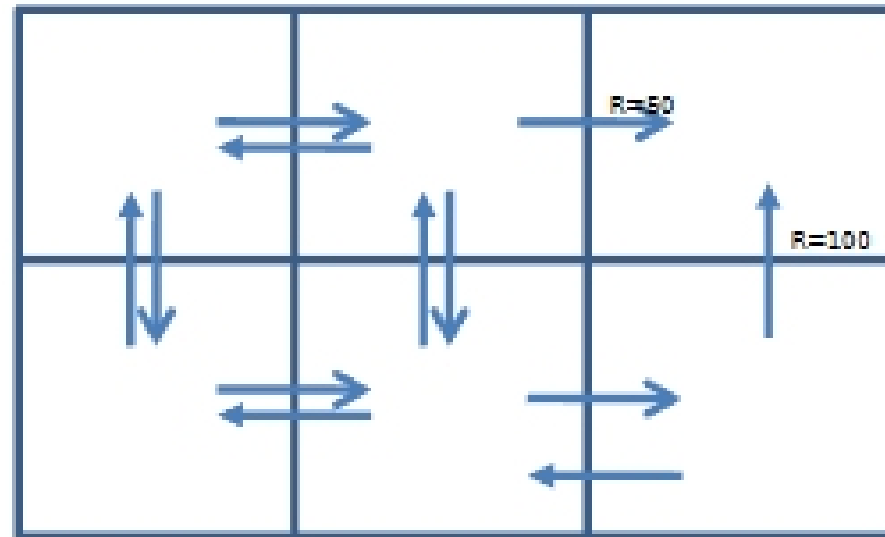
Reinforcement Learning

Some slides taken from previous 10701 recitations/lectures

A (Fully Deterministic) World



World



A **policy** is a mapping from **State** => **Action**. Normally denoted as $\pi(x)=a$
“**What action do I make if I find myself in a particular place?**”

Possible Questions.

- 1. If I am in state X. What is the value of following a particular policy?**
- 2: What is the best policy?**