CS 440
Intro to Artificial Intelligence

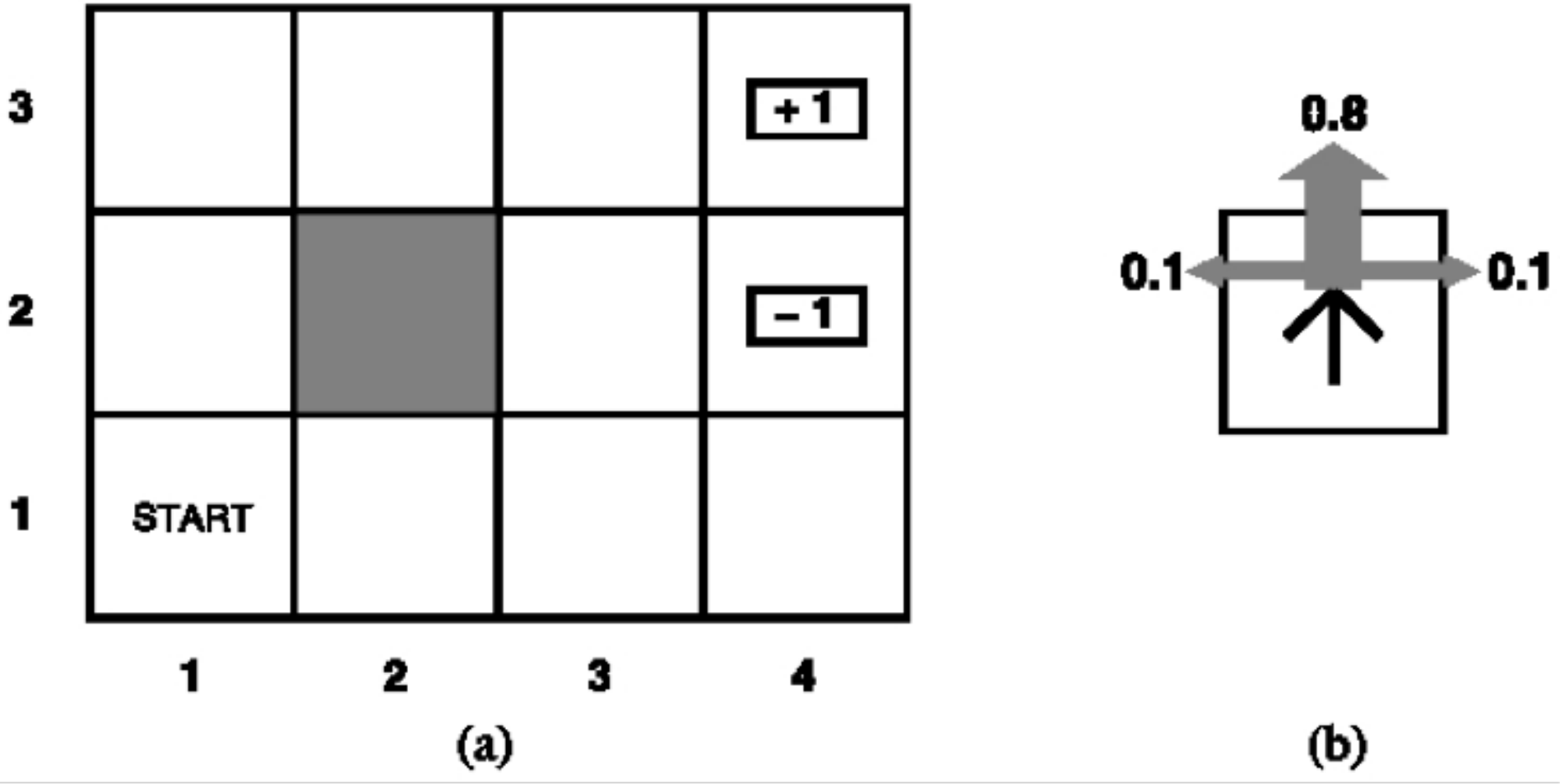
Lecture 20:

Sequential Decision Making -

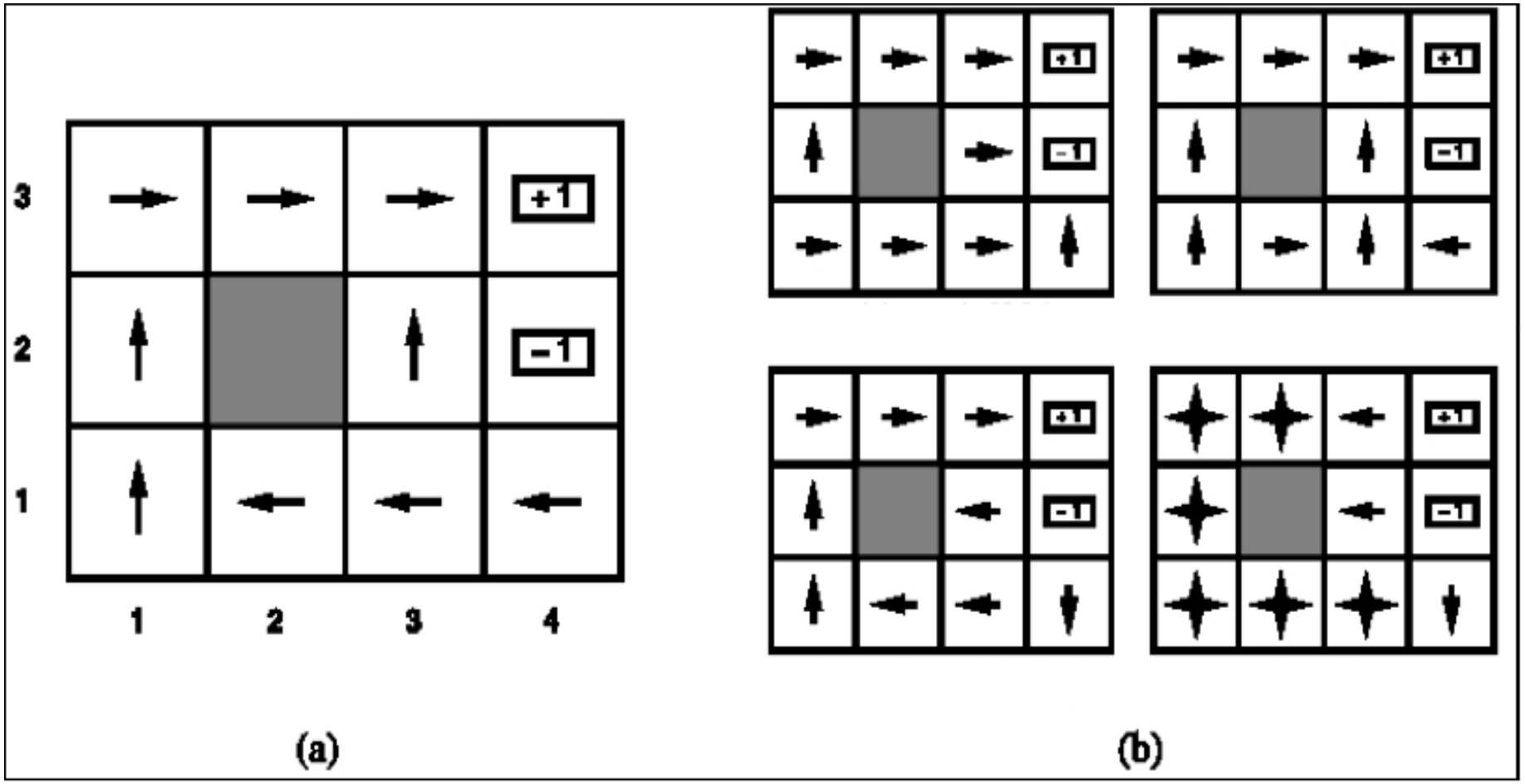
Partially Observable Markov Decision Processes

April 2, 2020

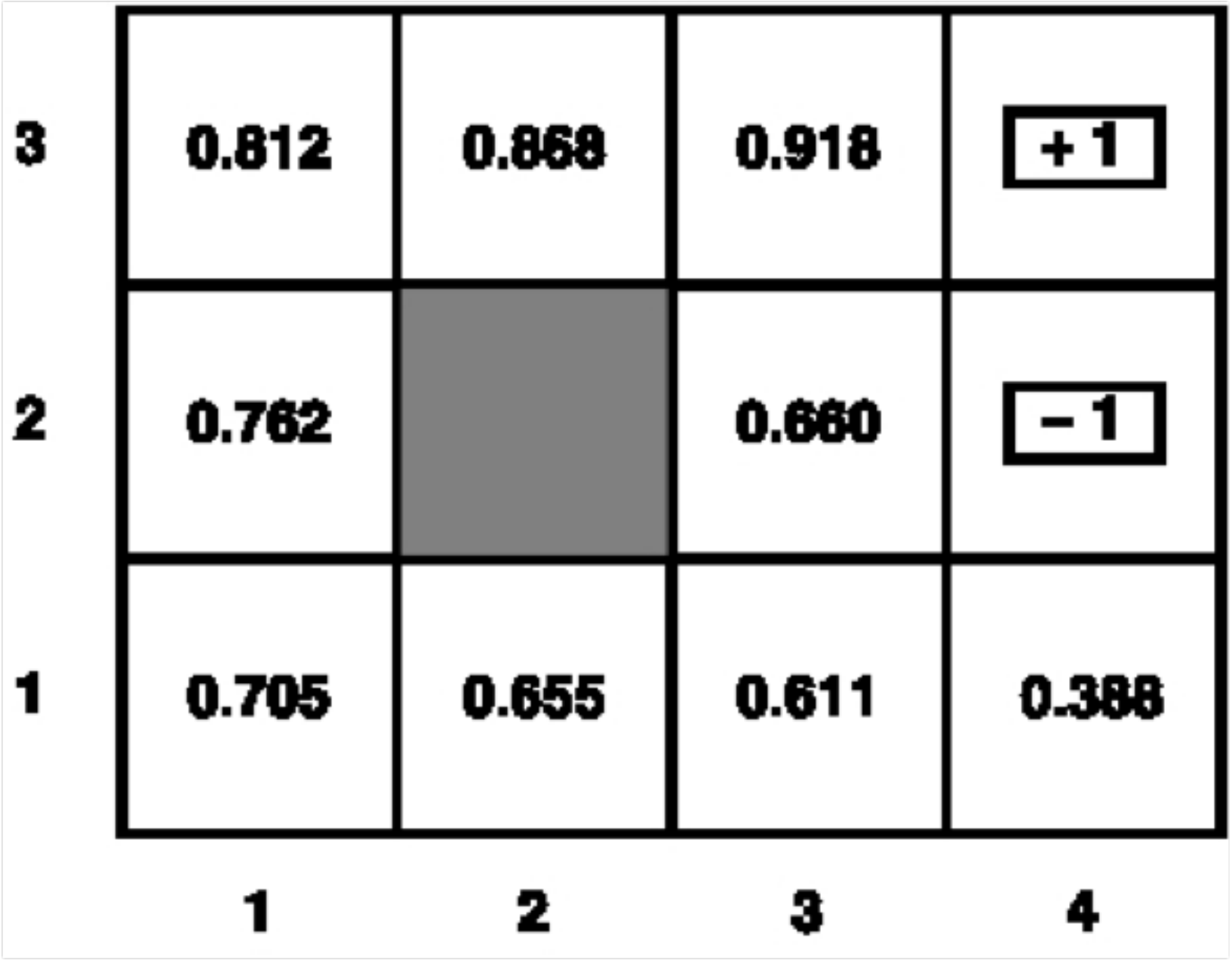
Sequential Decision Making



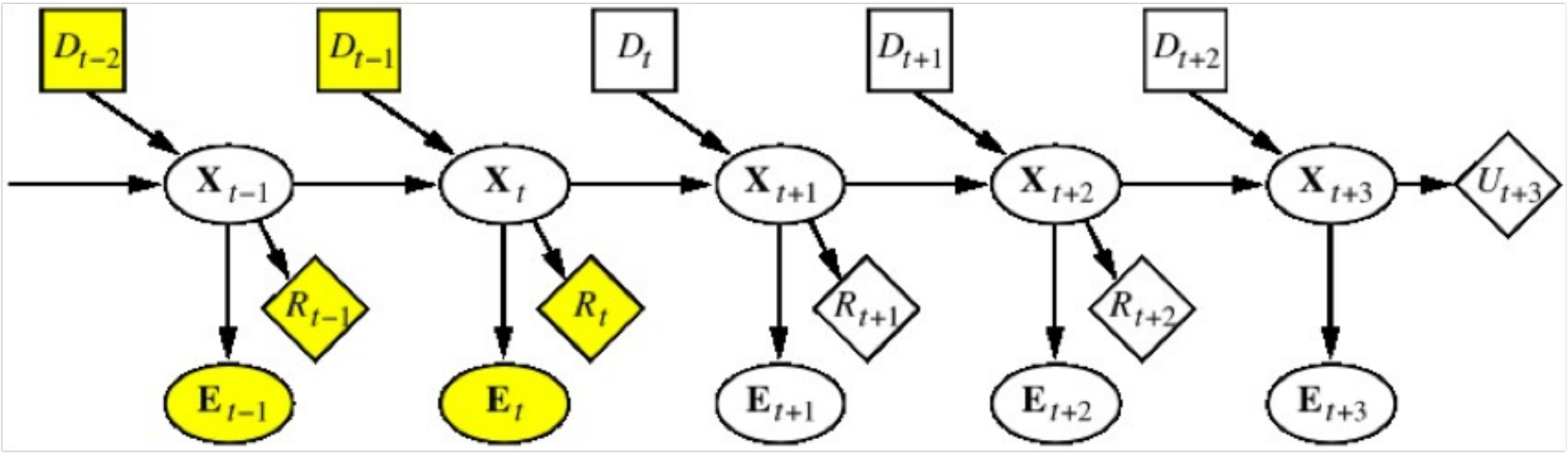
Optimal Policies



Utilities of States



Decision Theoretic Agents



Decision Theoretic Agents

D_t in $\mathbf{P}(\mathbf{X}_t \mid \mathbf{E}_{1:t})$

$\mathbf{E}_{t+1}$

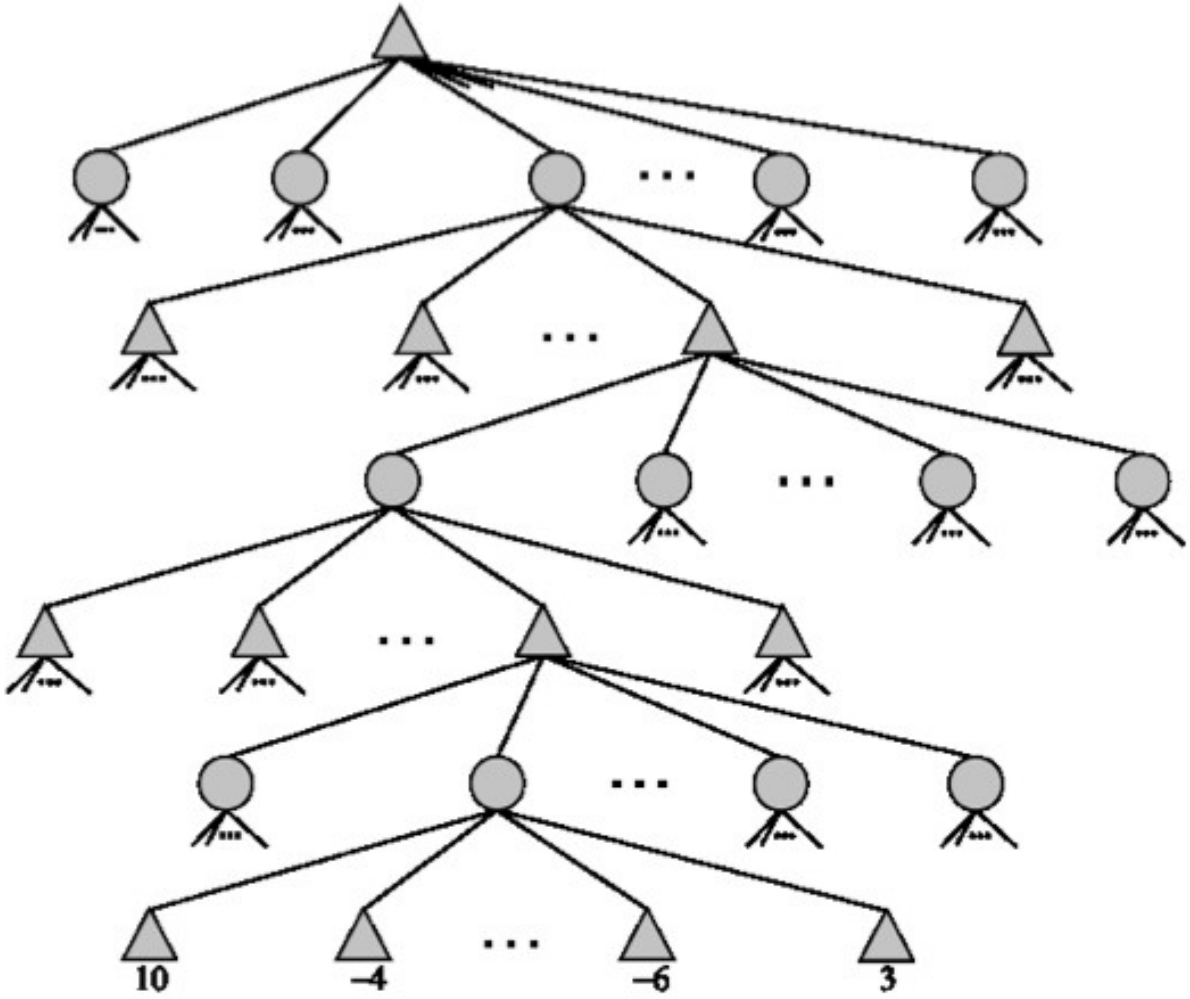
D_{t+1} in $\mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{E}_{1:t+1})$

$\mathbf{E}_{t+2}$

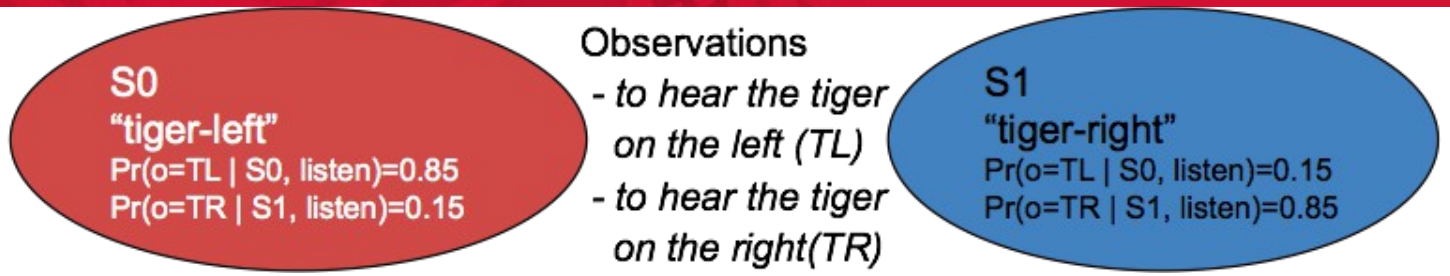
D_{t+2} in $\mathbf{P}(\mathbf{X}_{t+2} \mid \mathbf{E}_{1:t+2})$

$\mathbf{E}_{t+3}$

$U(\mathbf{X}_{t+3})$



POMDP example: The tiger problem



Prob. (LISTEN)	Tiger: left	Tiger: right
Tiger: left	1. 0	0. 0
Tiger: right	0. 0	1. 0

Listening does not change the position of the tiger

Reward Function
- *Penalty for wrong opening: -100*
- *Reward for correct opening: +10*
- *Cost for listening action: -1*

Prob. (LEFT)	Tiger: left	Tiger: right
Tiger: left	0. 5	0. 5
Tiger: right	0. 5	0. 5

The position of the tiger resets after we open one of the two doors

Prob. (RIGHT)	Tiger: left	Tiger: right
Tiger: left	0. 5	0. 5
Tiger: right	0. 5	0. 5