CS325 Artificial Intelligence Markov Decision Process Review

Cengiz Günay, Emory Univ.

(AN UNMATCHED LEFT PARENTHESIS OREATES AN UNRESOLVED TENSION THAT WILL STAY WITH YOU ALL DAY.

Spring 2013

Günay

Markov Decision Process Review

## Midterm Over, Let's Review Some and Have Fun

#### CajunBot I and II:





$$V(s) \leftarrow \left[ \arg \max_{a} \gamma \sum_{s'} P(s'|a) V(s') \right] + R(s).$$

$$1 \quad 2 \quad 3 \quad 4$$

$$a \quad b \quad -100 \quad +100$$

$$V(s) \leftarrow \left[ \arg \max_{a} \gamma \sum_{s'} P(s'|a) V(s') \right] + R(s).$$



Movement probability:

• *p* succeeds

• 
$$(1-p)$$
 fails, go reverse

$$V(s) \leftarrow \left[ \arg \max_{a} \gamma \sum_{s'} P(s'|a) V(s') \right] + R(s).$$

$$1 \quad 2 \quad 3 \quad 4 \qquad \text{Movement probability:}$$

$$a \quad b \quad -100 \qquad +100 \qquad \bullet \ p \text{ succeeds}$$

$$\bullet \quad (1-p) \text{ fails, go reverse}$$

Question: Find final values of all positions.

۲

Günay

$$V(s) \leftarrow \left[ \arg \max_{a} \gamma \sum_{s'} P(s'|a) V(s') \right] + R(s).$$

$$1 \quad 2 \quad 3 \quad 4 \qquad \text{Movement probability:}$$

$$a \quad b \quad -100 \qquad +100 \qquad \bullet \quad p \text{ succeeds}$$

$$\bullet \quad (1-p) \text{ fails, go reverse}$$

Question: Find final values of all positions.

• For 
$$p = 1$$
,  $\gamma = 1$ ,  $\text{cost} = -4$ 

$$V(s) \leftarrow \left[ rg\max_{a} \gamma \sum_{s'} P(s'|a) V(s') 
ight] + R(s).$$

	1	2	3	4
а	84	88	92	96
b	-100	92	96	+100

Movement probability:

• p succeeds

• 
$$(1 - p)$$
 fails, go reverse

Question: Find final values of all positions.

• For 
$$p = 1$$
,  $\gamma = 1$ ,  $\text{cost} = -4$ 

$$V(s) \leftarrow \begin{bmatrix} \arg\max_{a} \gamma \sum_{s'} P(s'|a)V(s') \end{bmatrix} + R(s).$$

$$1 \quad 2 \quad 3 \quad 4 \qquad \text{Movement probability:}$$

$$a \quad -100 \qquad +100 \qquad e \quad p \text{ succeeds}$$

$$(1-p) \text{ fails, go reverse}$$

Question: Find final values of all positions.

• For 
$$p = 0.8$$
,  $\gamma = 1$ ,  $cost = -4$ 

• Value of a4 after one iteration?

$$V(s) \leftarrow \left[ rg\max_{a} \gamma \sum_{s'} P(s'|a) V(s') 
ight] + R(s).$$



Movement probability:

• p succeeds

• 
$$(1 - p)$$
 fails, go reverse

Question: Find final values of all positions.

• For 
$$p = 0.8$$
,  $\gamma = 1$ ,  $cost = -4$ 

• Value of a4 after one iteration?

$$V(s) \leftarrow \begin{bmatrix} \arg \max_{a} \gamma \sum_{s'} P(s'|a) V(s') \end{bmatrix} + R(s).$$

$$1 \quad 2 \quad 3 \quad 4 \qquad \text{Movement probability:}$$

$$-100 \quad +100 \qquad \bullet \ p \text{ succeeds}$$

$$\bullet \ (1-p) \text{ fails, go reverse}$$

Question: Find final values of all positions.

• For 
$$p = 0.8$$
,  $\gamma = 1$ ,  $cost = -4$ 

• Value of a4 after it converges?

a b

$$V(s) \leftarrow \left[ \arg \max_{a} \gamma \sum_{s'} P(s'|a) V(s') \right] + R(s).$$



Movement probability:

• *p* succeeds

• 
$$(1-p)$$
 fails, go reverse

Question: Find final values of all positions.

• For 
$$p = 0.8$$
,  $\gamma = 1$ ,  $cost = -4$ 

• Value of a4 after it converges?

$$V(s) \leftarrow \left[ \arg \max_{a} \gamma \sum_{s'} P(s'|a) V(s') \right] + R(s).$$

$$1 \quad 2 \quad 3 \quad 4 \qquad \text{Movement probability:}$$

$$p \text{ succeeds}$$

$$(s = a + 100) \quad (s = b + 100)$$

а b -100+100

• 
$$(1 - p)$$
 fails, go reverse

Question: Find final values of all positions.

• For 
$$p = 0.8$$
,  $\gamma = 1$ ,  $cost = -4$ 

• Optimal policy after convergence?

$$V(s) \leftarrow \begin{bmatrix} \arg\max_{a} \gamma \sum_{s'} P(s'|a) V(s') \end{bmatrix} + R(s).$$

$$1 \qquad 2 \qquad 3 \qquad 4 \qquad \text{Movement probability:} \\ a \qquad \rightarrow \rightarrow \downarrow \qquad \downarrow \qquad \qquad \bullet \qquad p \text{ succeeds} \\ b \qquad -100 \qquad \downarrow \qquad \rightarrow \qquad +100 \qquad \qquad \bullet \qquad (1-p) \text{ fails, go reverse} \end{bmatrix}$$

Question: Find final values of all positions.

• For 
$$p = 0.8$$
,  $\gamma = 1$ ,  $cost = -4$ 

• Optimal policy after convergence?

$$V(s) \leftarrow \left[ \arg \max_{a} \gamma \sum_{s'} P(s'|a) V(s') \right] + R(s), \quad P(s,a) = 0.8$$



Günay

$$V(s) \leftarrow \left[ \arg \max_{a} \gamma \sum_{s'} P(s'|a) V(s') \right] + R(s), \quad P(s,a) = 0.8$$

$$\begin{array}{c} 85 & 83 & 93 + 100 \\ \hline 81 & 68 - 100 \\ \hline 77 & 73 & 70 & 47 \end{array} \qquad \begin{array}{c} y = 1 \\ R = -3 \\ \hline \hline 77 & 73 & 70 & 47 \end{array}$$

Günay

Markov Decision Process Review

$$V(s) \leftarrow \left[ \arg \max_{a} \gamma \sum_{s'} P(s'|a) V(s') \right] + R(s), \quad P(s,a) = 0.8$$

$$V(s) \leftarrow \left[ \arg \max_{a} \gamma \sum_{s'} P(s'|a) V(s') \right] + R(s), \quad P(s,a) = 0.8$$

$$V(s) \leftarrow \left[ \arg \max_{a} \gamma \sum_{s'} P(s'|a) V(s') \right] + R(s), \quad P(s,a) = 0.8$$

Günay

$$V(s) \leftarrow \left[ \arg\max_{a} \gamma \sum_{s'} P(s'|a) V(s') \right] + R(s), \quad P(s,a) = 0.8$$

$$V(s) \leftarrow \left[ \arg \max_{a} \gamma \sum_{s'} P(s'|a) V(s') \right] + R(s), \quad P(s,a) = 0.8$$

# Exploration and Prediction with MDPs and Reinforcement Learning





イロト イポト イヨト イヨト

## Midterm Evaluations

Please post your evaluations! You have the chance to grade me! :)

## Midterm Evaluations

Please post your evaluations! You have the chance to grade me! :)

• Any questions about midterm problems?

#### Midterm Evaluations

Please post your evaluations! You have the chance to grade me! :)

- Any questions about midterm problems?
- Have a Happy Spring Break!

