

# CS 4100/5100: Foundations of AI

## MDPs/RL

Instructor: Rob Platt  
[rplatt@ccs.neu.edu](mailto:rplatt@ccs.neu.edu)

College of Computer and information Science  
Northeastern University

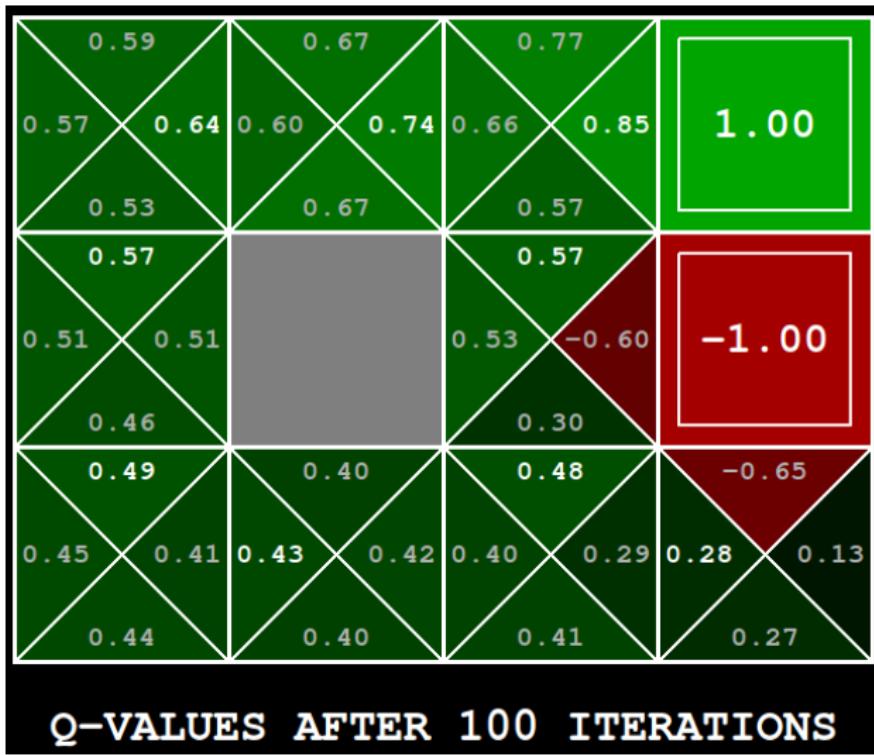
Fall, 2014

## Gridworld values



- ▶ process noise: 0.2 chance of doing something other than what was expected.
- ▶  $\gamma = 0.9$

# Gridworld Q values



## Value iteration example

▲ 0.00	▲ 0.00	▲ 0.00	0.00
▲ 0.00		▲ 0.00	0.00
▲ 0.00	▲ 0.00	▲ 0.00	0.00

**VALUES AFTER 0 ITERATIONS**

## Value iteration example

0.00	0.00	0.00	1.00
0.00		0.00	-1.00
0.00	0.00	0.00	0.00

**VALUES AFTER 1 ITERATIONS**

## Value iteration example

0.00	0.00	0.72	1.00
0.00		0.00	-1.00
0.00	0.00	0.00	0.00

**VALUES AFTER 2 ITERATIONS**

## Value iteration example

0.00 ↗	0.52 ↗	0.78 ↗	1.00
↑ 0.00		↑ 0.43	-1.00
↑ 0.00	↑ 0.00	↑ 0.00	0.00
			↓ 0.00

**VALUES AFTER 3 ITERATIONS**

## Value iteration example

0.37 ▶	0.66 ▶	0.83 ▶	1.00
▲ 0.00		▲ 0.51	-1.00
▲ 0.00	0.00 ▶	▲ 0.31	◀ 0.00
VALUES AFTER 4 ITERATIONS			

## Value iteration example

0.51 ↗	0.72 ↗	0.84 ↗	1.00
0.27 ↑		0.55 ↑	-1.00
0.00 ↑	0.22 ↗	0.37 ↑	0.13 ←
VALUES AFTER 5 ITERATIONS			

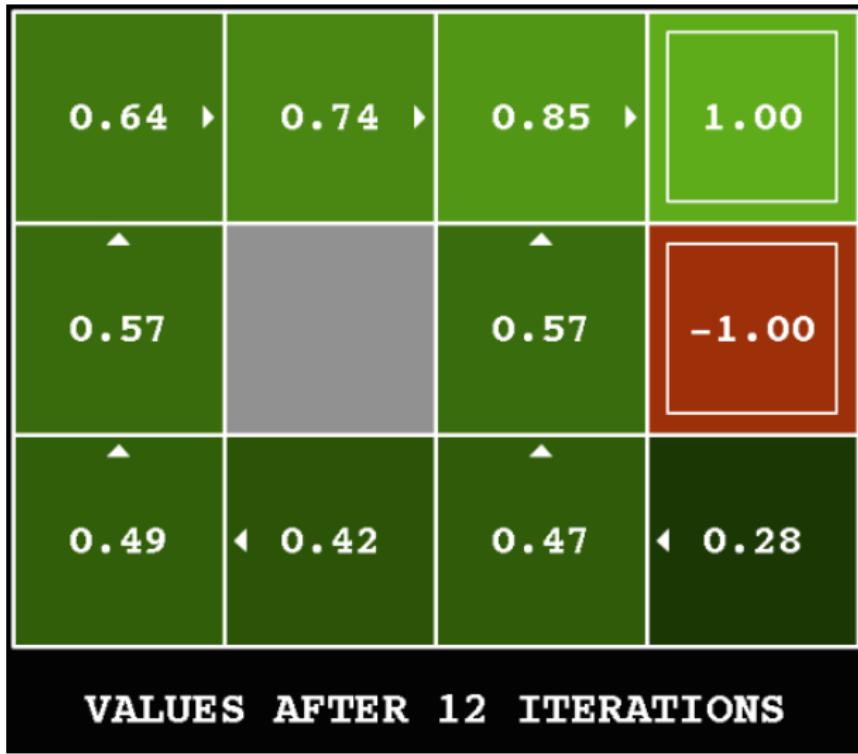
## Value iteration example

0.62 ↗	0.74 ↗	0.85 ↗	1.00
0.50 ↑		0.57 ↑	-1.00
0.34 ↑	0.36 ↗	0.45 ↑	0.24 ←
VALUES AFTER 7 ITERATIONS			

## Value iteration example

0.64 ▶	0.74 ▶	0.85 ▶	1.00
▲ 0.55		▲ 0.57	-1.00
▲ 0.46	0.40 ▶	0.47	◀ 0.27
VALUES AFTER 9 ITERATIONS			

## Value iteration example



## Value iteration example

0.64 →	0.74 →	0.85 →	1.00
▲ 0.57		▲ 0.57	-1.00
▲ 0.49	◀ 0.43	▲ 0.48	◀ 0.28
VALUES AFTER 100 ITERATIONS			