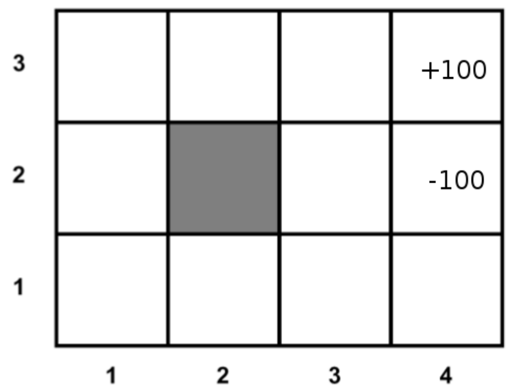
HW3: RL and Probability

CS4300: Artificial Intelligence **Alan Kuntz**
University of Utah

1 Reinforcement Learning

Consider the following gridworld:



Suppose that we run two episodes that yield the following sequences of (state, action, reward) tuples:

S	A	R	S	A	R
(1,1)	right	-1	(1,1)	up	-1
(2,1)	left	-1	(1,2)	up	-1
(1,1)	up	-1	(1,3)	right	-1
(1,2)	up	-1	(2,3)	right	-1
(1,3)	up	-1	(3,3)	right	-1
(2,3)	right	-1	(4,3)	exit	+100
(3,3)	right	-1	(done)		
(4,3)	exit	+100			
(done)					

1. According to direct estimation, what are the values for every state in the grid?
2. According to model-based learning, what are the transition probabilities for every (state, action, state) triple. Don't bother listing all the ones that we have no information about.
3. Suppose that we run Q-learning. However, instead of initializing all our Q values to zero, we initialize them to some large positive number ("large" with respect to the maximum reward possible in the world: say, 10 times the max reward). I claim that this will cause a Q-learning agent to initially explore a lot and then eventually start exploiting. Why should this be true? Justify your answer in a short paragraph.

2 Joint Probabilities

Given the following joint distribution table:

Weather	Temperature	Probability
snow	cold	0.1
snow	warm	0.4
sun	cold	0.3
sun	warm	0.2

Answer the following and show your work:

1. What are the random variables?
2. What is the domain of Temperature?
3. What is P(Weather = snow, Temperature = warm)?
4. What is P(sun)?
5. What is the probability that it is snowing OR it is cold?

3 Conditional Probabilities

Given the following joint distribution table:

X	Y	P
+x	+y	0.4
+x	-y	0.1
-x	+y	0.3
-x	-y	0.2

Answer the following and show your work:

1. What is $P(+x|+y)$?
2. What is $P(-x|+y)$?
3. What is $P(+x|-y)$?
4. What is $P(-x|-y)$?
5. What is $P(-y|+x)$?
6. What is $P(Y|+x)$?

4 No one warned me about the earthquakes...

Dr. Kuntz's dog barks a lot. He (the dog, not Dr. Kuntz) barks for many reasons. In the last year, we've observed him bark under three events (Event = m,q,n): when the mailperson is here (m), when an earthquake happens (q), and when neither is happening (n). We have the probability distribution over the events as:

Event	P
m	0.1
q	0.01
n	0.89

and the conditional probability distributions of barking given an event as P(B = {+b,-b}|Event):

Barking	Event	P
+b	m	0.8
-b	m	0.2
+b	q	0.9
-b	q	0.1
+b	n	0.4
-b	n	0.6

Note that this is 3 conditional distributions listed in one table, similar to as in the lectures.

Oh no! Dr. Kuntz's dog is barking again! We want to know the probability that an event is happening given the evidence of the dog barking.

Given that the dog is barking please compute the probability of each type of event. In other words, what is P(Event|+b)? Show your work.

Please tell me soon, I need to know whether to get the mail, to take cover, or to just ignore the dog.