CSCI 561 - Foundation for Artificial Intelligence

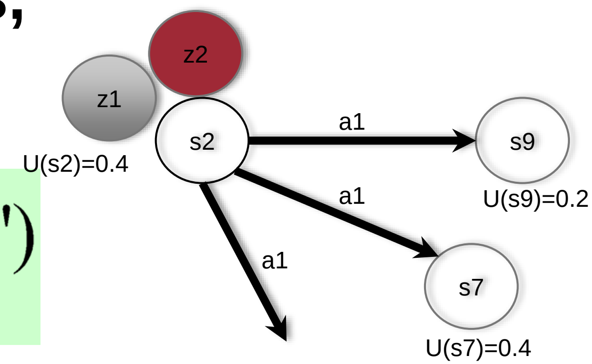
Discussion Section (Week 14)
MDP and Bayesian Learner

PROF WEI-MIN SHEN WMSHEN@USC.EDU

Maximum Expected Utility (MEU) and Rational Agents

- Every state has a utility value $U(s)$
- The expected utility of an action given the current evidence or observation e, is the average utility value of the outcomes, weighted by the probability that the outcome occurs:

$$EU(a|e) = \sum_{s'} P(result(a) = s' | a, e) U(s')$$



- The principle of maximum expected utility (MEU) is that a rational agent should choose the action that maximizes its expected utility:

$$action = \arg\max_{a} EU(a|e)$$

16.15 Making Simple Decisions

Consider a student who has the choice to buy or not buy a textbook for a course. We'll model this as a decision problem with one Boolean decision node, B, indicating whether the agent chooses to buy the book, and two Boolean chance nodes, M, indicating whether the student has mastered the material in the book, and P , indicating whether the student passes the course. Of course, there is also a utility node, U. A certain student, Sam, has an additive utility function: 0 for not buying the book and -$100 for buying it; and $2000 for passing the course and 0 for not passing. Sam's conditional probability estimates are as follows:

$P(p|b,m) = 0.9$
$P(m|b) = 0.9$
$P(p|b, \neg m) = 0.5$
$P(m|\neg b) = 0.7$
$P(p|\neg b, m) = 0.8$
$P(p|\neg b, \neg m) = 0.3$

You might think that P would be independent of B given M, But this course has an open-book final—so having the book helps.

Draw the decision network for this problem.

P(p|b,m) = 0.9
P(m|b) = 0.9
P (p|b, ¬m) = 0.5
P (m|¬b) = 0.7
P(p|¬b, m) = 0.8
P (p|¬b, ¬m) = 0.3

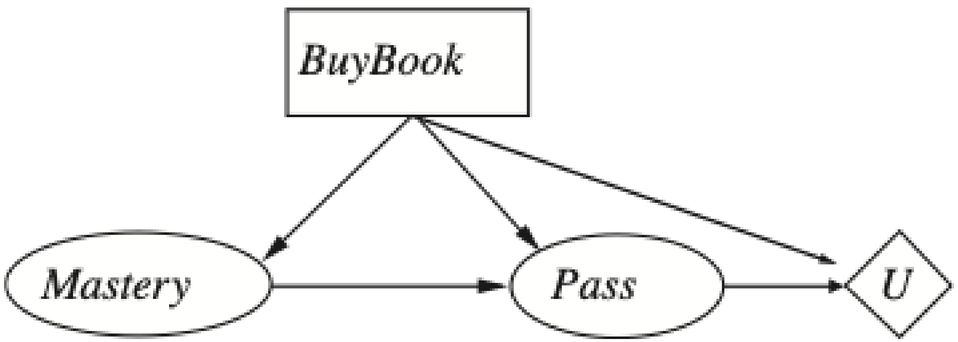


Figure S16.1 A decision network for the book-buying problem.

Utility	Buy	~Buy
Pass	2000-100	2000
~Pass	0-100	0+0

Compute the expected utility of buying the book and of not buying it.

For each of $B = b$ and $B = \neg b$, we compute $P(p|B)$ and thus $P(\neg p|B)$ by marginalizing out M, then use this to compute the expected utility.

$$
\begin{aligned}
P(p|b) &= \sum_m P(p|b, m)P(m|b) \\
&= 0.9 \times 0.9 + 0.5 \times 0.1 \\
&= 0.86 \\
P(p|\neg b) &= \sum_m P(p|\neg b, m)P(m|\neg b) \\
&= 0.8 \times 0.7 + 0.3 \times 0.3 \\
&= 0.65
\end{aligned}
$$

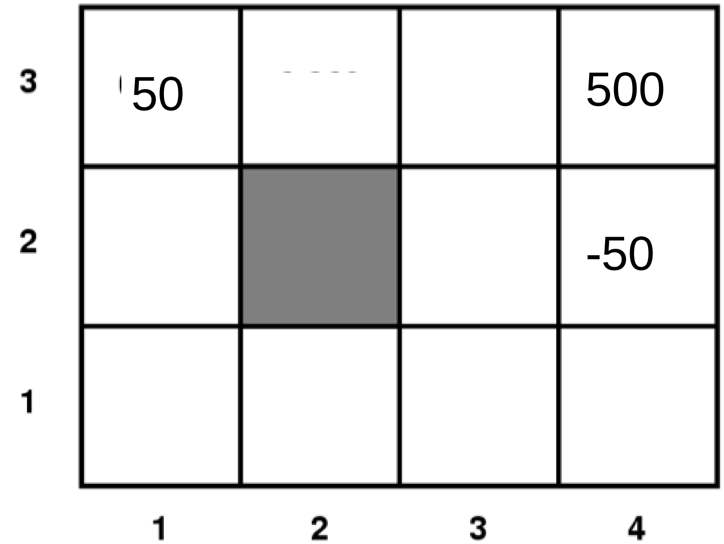
Compute the expected utility of buying the book and of not buying it.

The expected utilities are thus:

$$
\begin{aligned}
EU[b] &= \sum_p P(p|b)U(p,b) \\
&= 0.86(2000 - 100) + 0.14(-100) \\
&= 1620 \\
EU[\neg b] &= \sum_p P(p|\neg b)U(p,\neg b) \\
&= 0.65 \times 2000 + 0.35 \text{ x } 0 \\
&= 1300
\end{aligned}
$$

BUY!

MDP



Given a Gridworld domain, where terminal states (1,3), (4,3), and (4,2) have rewards 50, 500, and -50 respectively, the set of possible actions are {N,E,S,W, or X for terminal states}, the agent moves deterministically, all V and Q values for non terminal states have been initialized to 0, answer the questions below.

Circle the letter that corresponds to the best answer for the question.

What are the optimal values, $V*$ of each state in the above grid if $\gamma = 0.5$, $c(a)=0$, $R(s)=0$ for non terminal states?

(Remember $V_{t+1}(s) = R(s) + \text{Max}_{a \varepsilon A} \{c(a)+\gamma \Sigma_{s' \varepsilon S} P(s'|a,s) V_t(s')\}$)

a. $V_{(1,1)}=15.75$, $V_{(1,2)}=25$, $V_{(2,1)}=31.25$, $V_{(2,3)}=125$, $V_{(3,1)}=62.5$, $V_{(3,2)}=125$, $V_{(3,3)}=250$, $V_{(4,1)}=25$
b. $V_{(1,1)}=12.5$, $V_{(1,2)}=25$, $V_{(2,1)}=31.25$, $V_{(2,3)}=125$, $V_{(3,1)}=62.5$, $V_{(3,2)}=125$, $V_{(3,3)}=250$, $V_{(4,1)}=31.25$
c. $V_{(1,1)}=15.625$, $V_{(1,2)}=25$, $V_{(2,1)}=31.25$, $V_{(2,3)}=125$, $V_{(3,1)}=62.5$, $V_{(3,2)}=125$, $V_{(3,3)}=250$, $V_{(4,1)}=31.25$
d. $V_{(1,1)}=12.5$, $V_{(1,2)}=25$, $V_{(2,1)}=25$, $V_{(2,3)}=25$, $V_{(3,1)}=50$, $V_{(3,2)}=100$, $V_{(3,3)}=250$, $V_{(4,1)}=25$
e. None of the above

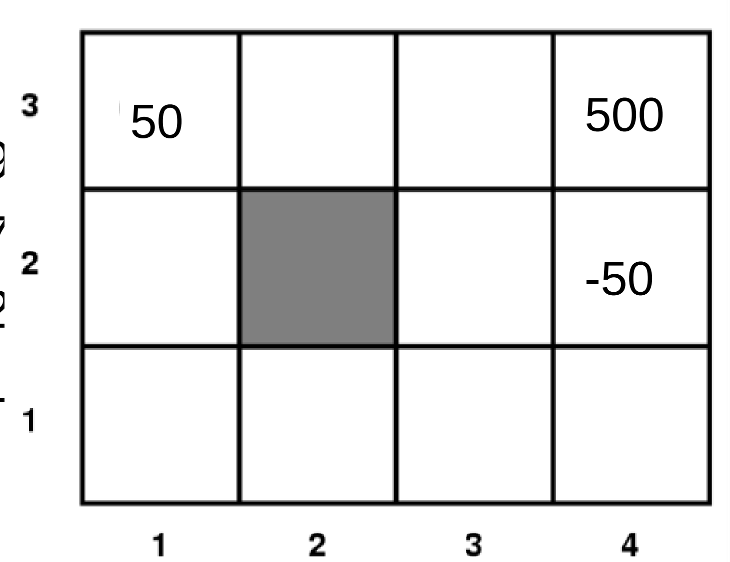
C

MDP

What are the Q values of state (3,2) in the above grid if γ = 0.5, c(a)=0, R(s)=-2 for non terminal states?

(Remember $Q_{t+1}(a,s) = R(s) + c(a) + \gamma \Sigma_{s' \varepsilon S} P(s'|a,s) \max_{a' \varepsilon A} Q_t(a's')$)

a. $Q_{((3,2),N)}$=122, $Q_{((3,2),E)}$=-27, $Q_{((3,2),S)}$=59
b. $Q_{((3,2),N)}$=122, $Q_{((3,2),E)}$=-27, $Q_{((3,2),S)}$=27
c. $Q_{((3,2),N)}$=125, $Q_{((3,2),E)}$=-25, $Q_{((3,2),S)}$=62
d. $Q_{((3,2),N)}$=120, $Q_{((3,2),E)}$=-27, $Q_{((3,2),S)}$=31
e. None of the above

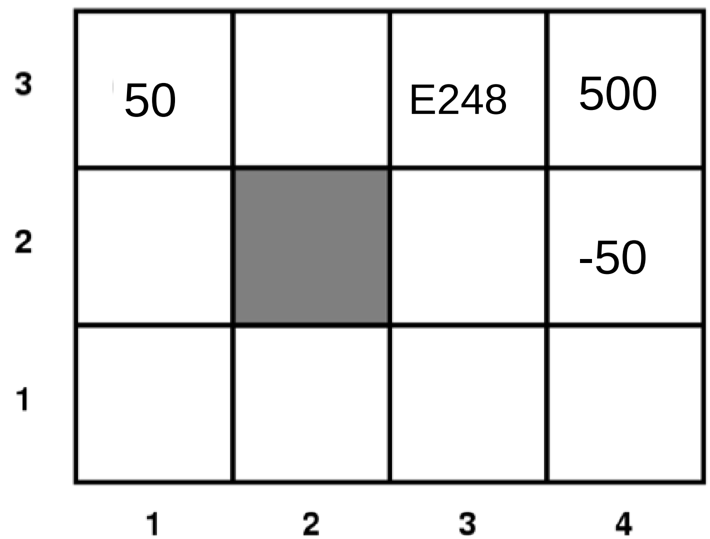


MDP

What are the Q values of state (3,2) in the above grid if $\gamma = 0.5$, $c(a)=0$, $R(s)=-2$ for non terminal states?

(Remember $Q_{t+1}(a,s) = R(s) + c(a)+\gamma\Sigma_{s'\varepsilon S} P(s'|a,s)\max_{a'\varepsilon A} Q_t(a's')$)

a. $Q_{((3,2),N)}=122$, $Q_{((3,2),E)}=-27$, $Q_{((3,2),S)}=5$
b. $Q_{((3,2),N)}=122$, $Q_{((3,2),E)}=-27$, $Q_{((3,2),S)}=2$
c. $Q_{((3,2),N)}=125$, $Q_{((3,2),E)}=-25$, $Q_{((3,2),S)}=6$
d. $Q_{((3,2),N)}=120$, $Q_{((3,2),E)}=-27$, $Q_{((3,2),S)}=3$
e. None of the above

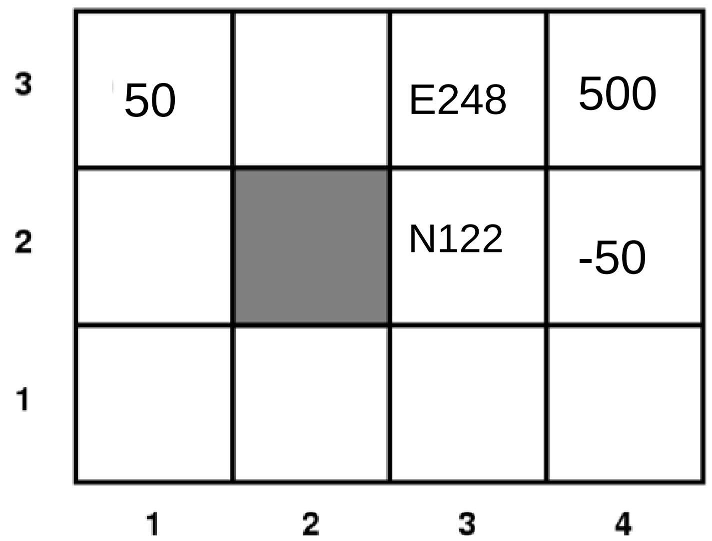


MDP

What are the Q values of state (3,2) in the above grid if γ = 0.5, c(a)=0, R(s)=-2 for non terminal states?

(Remember $Q_{t+1}(a,s) = R(s) + c(a) + \gamma \Sigma_{s' \varepsilon S} P(s'|a,s) \max_{a' \varepsilon A} Q_t(a's')$)

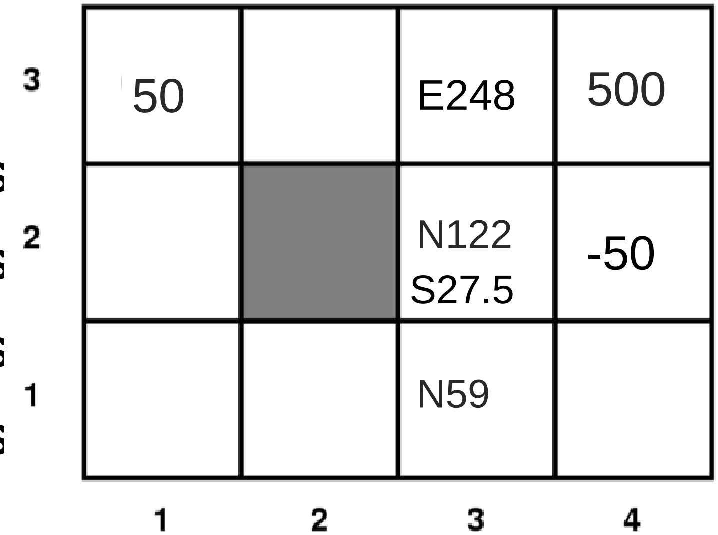
a. $Q_{((3,2),N)}$=122, $Q_{((3,2),E)}$=-27, $Q_{((3,2),S)}$=5
b. $Q_{((3,2),N)}$=122, $Q_{((3,2),E)}$=-27, $Q_{((3,2),S)}$=2
c. $Q_{((3,2),N)}$=125, $Q_{((3,2),E)}$=-25, $Q_{((3,2),S)}$=6
d. $Q_{((3,2),N)}$=120, $Q_{((3,2),E)}$=-27, $Q_{((3,2),S)}$=3
e. None of the above



MDP

What are the Q values of state (3,2) in the above grid if γ = 0.5, c(a)=0, R(s)=-2 for non terminal states?

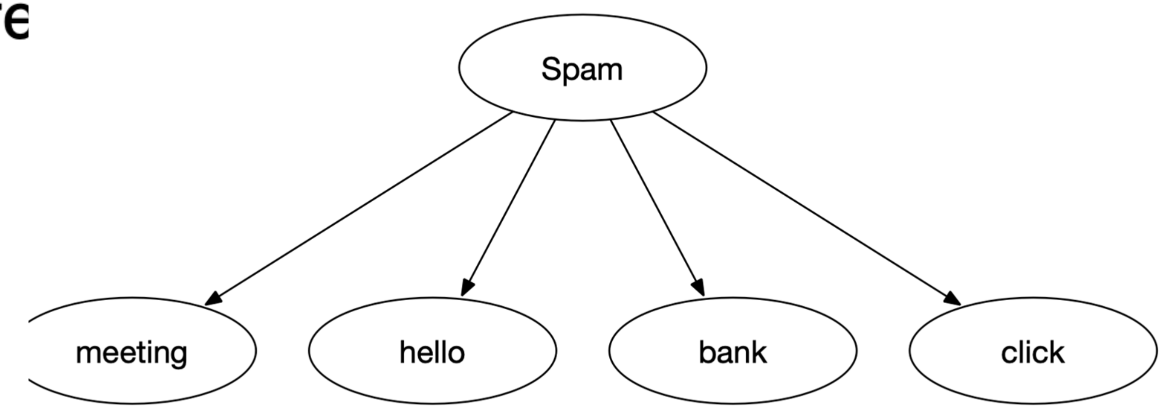
(Remember $Q_{t+1}(a,s) = R(s) + c(a)+\gamma\Sigma_{s'\varepsilon S} P(s'|a,s)\max_{a'\varepsilon A} Q_t(a's')$)



a. **$Q_{((3,2),N)}$=122, $Q_{((3,2),E)}$=-27, $Q_{((3,2),S}$**
b. **$Q_{((3,2),N)}$=122, $Q_{((3,2),E)}$=-27, $Q_{((3,2),S}$**
c. **$Q_{((3,2),N)}$=125, $Q_{((3,2),E)}$=-25, $Q_{((3,2),S}$**
d. **$Q_{((3,2),N)}$=120, $Q_{((3,2),E)}$=-27, $Q_{((3,2),S}$**
e. B **None of the above**

Naïve Bayes Classifier

- Naive Bayes classification for text categorization
  - Very common baseline (can do surprisingly well)
  - Classify news stories, ..
  - Classify spam vs. not-spam
  - Bag-of-words feature
  - P(Spam | words)



Naïve Bayes

- Suppose I want to know if a news article is about sports, politics, entertainment.
  - Classes: sports, politics, entertainment
- Probability that a document *d* belongs to class *c*
- Probability of class *c* given document *d*

$$P(c|d) = \frac{P(c)\,P(d|c)}{P(d)}$$

- Compute for every class!

SPAM	click for pharmacy
OK	free time today
SPAM	online pharmacy link
OK	no free time
OK	free good pharmacy
SPAM	pharmacy free link
OK	for time today
OK	time is money

$P(spam) =$

Maximum likelihood estimate

Vocabulary size: 12

click	online
for	link
pharmacy	no
free	good
time	is
today	money

$P(spam) = \frac{3}{8}$ $\quad P(\neg spam) = \frac{5}{8}$

SPAM	click for pharmacy
OK	free time today
SPAM	online pharmacy link
OK	no free time
OK	free good pharmacy
SPAM	pharmacy free link
OK	for time today
OK	time is money

$$P(spam) = \frac{3}{8} \qquad P(\neg spam) = \frac{5}{8}$$

$$P(pharmacy|spam) = 1/3$$

$$P(pharmacy|\neg spam) = 1/15$$

Msg = “Pharmacy for pharmacy”

$$P(spam|Msg) = \frac{P(spam)P(Msg|spam)}{P(spam)P(Msg|spam) + P(\neg spam)P(Msg|\neg spam)}$$

$$P(Msg|spam) = P(w_1|spam)P(w_2|spam)P(w_3|spam)$$

SPAM	click for pharmacy
OK	free time today
SPAM	online pharmacy link
OK	no free time
OK	free good pharmacy
SPAM	pharmacy free link
OK	for time today
OK	time is money

$$P(spam) = \frac{3}{8} \qquad P(\neg spam) = \frac{5}{8}$$

$$P(pharmacy|spam) = 1/3$$

$$P(pharmacy|\neg spam) = 1/15$$

$$P(spam|Msg) = \frac{3/8 \cdot 1/3 \cdot 1/9 \cdot 1/3}{P(spam)P(Msg|spam) + P(\neg spam)P(Msg|\neg spam)}$$

SPAM	click for pharmacy
OK	free time today
SPAM	online pharmacy link
OK	no free time
OK	free good pharmacy
SPAM	pharmacy free link
OK	for time today
OK	time is money

$$P(spam) = \frac{3}{8} \qquad P(\neg spam) = \frac{5}{8}$$

$$P(pharmacy|spam) = 1/3$$

$$P(pharmacy|\neg spam) = 1/15$$

Msg = “Pharmacy for pharmacy”

$$P(spam|Msg) = \frac{1/216}{1/216 + P(\neg spam)P(Msg|\neg spam)}$$

$$P(spam|Msg) = \frac{1/216}{1/216 + 5/8 \cdot 1/15 \cdot 1/15 \cdot 1/15}$$

SPAM	click for pharmacy
OK	free time today
SPAM	online pharmacy link
OK	no free time
OK	free good pharmacy
SPAM	pharmacy free link
OK	for time today
OK	time is money

$$P(spam) = \frac{3}{8} \qquad P(\neg spam) = \frac{5}{8}$$

$$P(pharmacy|spam) = 1/3$$

$$P(pharmacy|\neg spam) = 1/15$$

Msg = “Pharmacy for pharmacy”

$$P(spam|Msg) = \frac{P(spam)P(Msg|spam)}{P(spam)P(Msg|spam) + P(\neg spam)P(Msg|\neg spam)}$$

$$P(Msg|spam) = P(w_1|spam)P(w_2|spam)P(w_3|spam)$$

$$P(spam|Msg) = 25/26 = 0.96$$

SPAM	click for pharmacy
OK	free time today
SPAM	online pharmacy link
OK	no free time
OK	free good pharmacy
SPAM	pharmacy free link
OK	for time today
OK	time is money

$$P(spam) = \frac{3}{8} \qquad P(\neg spam) = \frac{5}{8}$$

$$P(pharmacy|spam) = 1/3$$

$$P(pharmacy|\neg spam) = 1/15$$

Msg = "Time for pharmacy"

$$P(spam|Msg) = \frac{P(spam)P(Msg|spam)}{P(spam)P(Msg|spam) + P(\neg spam)P(Msg|\neg spam)}$$

$$P(Msg|spam) = P(w_1|spam)P(w_2|spam)P(w_3|spam)$$

Want more?

Exercises 16.5, 17.1, 17.9, 17.10