1. (16 points) True/False

For the following questions, a correct answer is worth 2 points, no answer is worth 1 point, and an incorrect answer is worth 0 points. Circle *true* or *false* to indicate your answer.

a) (*true* or *false*) Inverse reinforcement learning (which makes helicopters fly by themselves) is primarily concerned with learning an expert's transition model.
Solution:
False. Inverse reinforcement learning focuses on learning an expert's *reward* function. Experts and lousy agents alike all use the same transition model.

b) (*true* or *false*) A* search will always expand fewer search nodes than uniform cost search.
Solution:
False. A heuristic can lead the search algorithm astray. Consider states S, 1, 2, 3 & G where S leads to 1 and 2 with cost 1, 1 leads to G with cost 1, and 2 leads to 3 with cost 2. UCS will expand S, 1, 2, G. For a heuristic that is zero everywhere except h(1) = 3, then A* will expand S, 2, 3, 1, G.

c) (*true* or *false*) K-means is a clustering algorithm that is guaranteed to converge.
Solution:
True. The convergence proof was given in class.

d) (*true* or *false*) A MIRA classifier is guaranteed to perform better on unseen data than a perceptron.
Solution:
False. MIRA provides no such guarantee.

e) (*true* or *false*) For a two-player zero-sum game tree of depth 4 or greater, alpha-beta pruning must prune at least one node.
Solution:
False. As a simple counterexample, if every player only has one legal move at each turn, then there will be no pruning.

f) (*true* or *false*) A CSP with only boolean variables can always be solved in polynomial time.
Solution:
False. For instance, a boolean CSP can encode a 3-SAT problem, which is NP-complete.

g) (*true* or *false*) Sampling from a Bayes net using likelihood weighting will systematically overestimate the posterior of a variable conditioned on one of its descendants.
Solution:
False. Likelihood weighting is an unbiased sampling procedure.

h) (*true* or *false*) Every discrete-valued dynamic Bayes net is equivalent to some hidden Markov model.
Solution:

True and *False* (everyone wins). The answer is *false* because dynamic Bayes nets can have dependence between evidence variables when conditioning on hidden variables, but a lecture slide says it's *true*. The general case for which DBNs are equivalent to HMMs is when no hidden variable is a child of an evidence variable in the DBN.

2. (14 points) MDP: Walk or Jump?

Consider an MDP with states $\{4, 3, 2, 1, 0\}$, where 4 is the starting state. In states $k \geq 1$, you can *walk* (W) and $T(k, W, k-1) = 1$. In states $k \geq 2$, you can also *jump* (J) and $T(k, J, k-2) = T(k, J, k) = 1/2$. State 0 is a terminal state. The reward $R(s, a, s') = (s - s')^2$ for all (s, a, s'). Use a discount of $\gamma = 1/2$.

(a) (3 pt) Compute $V^*(2)$.

Solution:

We compute the value function from state 0 to state 2.

$$\begin{aligned}
V^*(0) &= 0 \\
V^*(1) &= \max\{1 + \gamma V^*(0), \frac{1}{2}(1 + \gamma V^*(0)) + \frac{1}{2}\gamma V^*(1)\} \\
&= \max\{1, \frac{1}{2} + \frac{1}{4}V^*(1)\} \\
&= \max\{1, \frac{1}{2}\frac{4}{3}\} \\
&= 1 \\
V^*(2) &= \max\{1 + \gamma V^*(1), \frac{1}{2}(4 + \gamma V^*(0)) + \frac{1}{2}\gamma V^*(2)\} \\
&= \max\{\frac{3}{2}, 2 + \frac{1}{4}V^*(2)\} \\
&= \max\{\frac{3}{2}, 2\frac{4}{3}\} = \frac{8}{3}
\end{aligned}$$

where $\frac{1}{2}\frac{4}{3}$ comes from supposing that if $\frac{1}{2} + \frac{1}{4}V^*(1)$ were the maximum, then

$$\begin{aligned}
V^*(1) &= \frac{1}{2} + \frac{1}{4}V^*(1) \\
\frac{3}{4}V^*(1) &= \frac{1}{2} \\
V^*(1) &= \frac{1}{2}\frac{4}{3}
\end{aligned}$$

and similarly for $2\frac{4}{3}$.

(b) (3 pt) Compute $Q^*(4, W)$?

Solution:

$$\begin{aligned}
V^*(3) &= \frac{1}{2}(4 + \frac{1}{2}V^*(1)) + \frac{1}{2}\frac{1}{2}V^*(3) \\
&= \frac{9}{4} + \frac{1}{4}V^*3 \\
&= 3 \\
Q^*(4, W) &= 1 + \frac{1}{2}V^*(3) = \frac{5}{2}
\end{aligned}$$

(c) **(4 pt)** Now consider the same MDP, but with infinite states $\{4, 3, 2, 1, 0, -1, ...\}$ and no terminal states. Like before, $T(k, J, k-2) = T(k, J, k) = 1/2$ and $T(k, W, k-1) = 1$. $R(s, a, s') = (s - s')^2$. Compute $V^*(2)$.

Solution:

By symmetry, $V^*(s)$ is constant. This implies:

$$\begin{aligned} V^*(2) &= \max\{1 + \gamma V^*(1), \frac{1}{2}(4 + \gamma V^*(0)) + \frac{1}{2}\gamma V^*(2)\} \\ &= \max\{1 + \gamma V^*(2), \frac{1}{2}4 + \gamma V^*(2)\} \\ &= \frac{2}{1-\gamma} \\ &= 4 \end{aligned}$$

(d) **(4 pt)** In the infinite MDP above, an agent acts randomly at each time step. With probability p, it *walks*. With probability $1 - p$, it *jumps*. What is its expected utility in terms of p, starting in state 4?

Solution:

Here, the expected utility is the expected sum of discounted rewards. We have a modified Bellman equation for this stochastic policy:

$$V^p(4) = p(1 + 0.5V^p(3)) + (1-p) * (0.5 * (4 + 0.5 * V^p(2)) + 0.5 * (0.5 * V^p(4)))$$

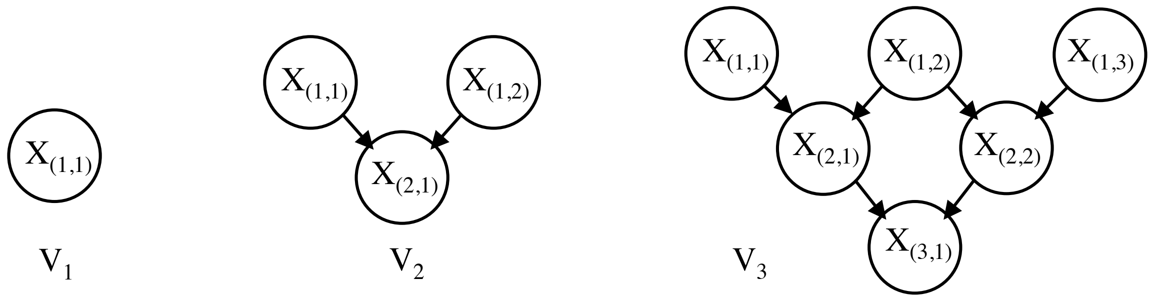
But again all states have the same value because of symmetry, so the above equation simplifies to:

$$V^p(4) = p(1 + 0.5V^p(4)) + (1-p) * (0.5 * 4 + 0.5 * V^p(4))$$

, so $V^p(4) = 4 - 2p$.

3. (21 points) Very Big V Structures

Let V_n be a Bayes net with nodes $X_{(i,j)}$ for all $i+j \leq n+1$, where both $i \geq 1$ and $j \geq 1$, and where the parents of $X_{(i,j)}$ are $X_{(i-1,j)}$ and $X_{(i-1,j+1)}$. Nodes $X_{(1,j)}$ have no parents. V_1, V_2, and V_3 are shown below.



(a) **(4 pt)** For any V_n, give general conditions in terms of i, j, k, and ℓ that guarantee $X_{(i,j)}$ independent of $X_{(k,\ell)}$ (i.e., $X_{(i,j)} \perp\!\!\!\perp X_{(k,\ell)}$), assuming $i < k$.

Solution:

In general, two nodes are independent (conditioned on nothing) if they are not both descendants of the same node, and neither is a descendent of the other. General conditions which enforce this are $l > j+i-1$ and $k+l < j+1$.

We also gave full credit to the nearly-correct answer of conditions that ensure the second is not a descendent of the first. $X_{(k,\ell)}$ is a descendent of $X_{(i,j)}$ if $k+\ell \geq i+j$ and $\ell \leq j$. So, $X_{(i,j)} \perp\!\!\!\perp X_{(k,\ell)}$ when $k+\ell < i+j$ or $\ell > j$.

Solutions that were too specific but otherwise correct received 1 point. Covering one of the two cases for either acceptable answer received 2 points.

(b) **(4 pt)** For V_n, give conditions in terms of i, j, k, and ℓ that guarantee $X_{(1,i)} \perp\!\!\!\perp X_{(1,j)} \mid X_{(k,\ell)}$ for $i < j$.

Solution:

$X_{(k,\ell)}$ must not be a descendent of both. So, either $k+\ell < j+1$, or $\ell > i$.

(c) **(2 pt)** When computing $P(X_{(3,1)})$ with variable elimination, what factor is created by first eliminating $X_{(1,2)}$ from V_3?

Solution:

$P(X_{(2,1)}, X_{(2,2)} | X_{(1,1)}, X_{(1,3)})$. Using factor notation $m(X_{(2,1)}, X_{(2,2)}, X_{(1,1)}, X_{(1,3)})$ or $m(X_{(2,1)}, X_{(2,2)} | X_{(1,1)}, X_{(1,3)})$ is also acceptable.

For the remaining parts of this question, consider V_4 where each $X_{(i,j)}$ takes values *true* and *false*.

(d) (2 pt) Given the factors below, fill in the table for $P(X_{(1,1)}|\neg x_{(3,1)})$, or state "not enough information."

$X_{(1,1)}$	$P(X_{(1,1)})$
$x_{(1,1)}$	1/3
$\neg x_{(1,1)}$	2/3

$X_{(1,1)}$	$X_{(3,1)}$	$P(X_{(3,1)}\|X_{(1,1)})$
$x_{(1,1)}$	$x_{(3,1)}$	1/3
$x_{(1,1)}$	$\neg x_{(3,1)}$	2/3
$\neg x_{(1,1)}$	$x_{(3,1)}$	0
$\neg x_{(1,1)}$	$\neg x_{(3,1)}$	1

$X_{(1,1)}$	$X_{(3,1)}$	$P(X_{(1,1)}\|\neg x_{(3,1)})$
$x_{(1,1)}$	$\neg x_{(3,1)}$	
$\neg x_{(1,1)}$	$\neg x_{(3,1)}$	

Solution:

Using Bayes' rule $P(x_{(1,1)}|\neg x_{(3,1)}) = 1/4$, $P(\neg x_{(1,1)}|\neg x_{(3,1)}) = 3/4$.

(e) (3 pt) In V_4, if $P(X_{(1,h)} = true) = \frac{1}{3}$ for all h, and the CPT for $X_{(i,j)}$ is defined below for all $i > 1$, what is the joint probability of all variables when $X_{(i,j)}$ is *true* if and only if $i + j$ is even.

$X_{(i,j)}$	$X_{(i-1,j)}$	$X_{(i-1,j+1)}$	$P(X_{(i,j)}\|X_{(i-1,j)}, X_{(i-1,j+1)})$
true	*true*	*true*	1
true	*false*	*true*	1/2
true	*true*	*false*	1/2
true	*false*	*false*	1/3

Clarification: Compute $P(X_{(1,1)} = true, X_{(1,2)} = false, X_{(1,3)} = true, ..., X_{(4,1)} = false)$

Solution:

$(\frac{1}{3})^2 \cdot (\frac{2}{3})^2 \cdot (\frac{1}{2})^6 = \frac{1}{1296}$

(f) (4 pt) Given the definition of V_4 above, formulate a CSP over variables $X_{(i,j)}$ that is satisfied by only and all assignments with non-zero joint probability.

Solution:

Variables: $X_{(i,j)}$; Values: *true*, *false*; Constraint: $(X_{(i,j)}, X_{(i-1,j)}, X_{(i-1,j+1)}) \neq$ (*false*, *true*, *true*) for all $i > 1$. Only the constraint is necessary for full credit, and the $i > 1$ condition is optional.

(g) (2 pt) Does there exist a binary CSP that is equivalent to the CSP you defined?

Solution:

Yes. All finite-valued CSPs are equivalent to some binary CSP.

4. (10 points) Linear Naive Bayes

Recall that a naive Bayes classifier with features F_i $(i = 1, \ldots, n)$ and label Y uses the classification rule:

$$\arg\max_y P(y|f_1, \ldots, f_n) = \arg\max_y P(y) \prod_{i=1}^{n} P(f_i|y)$$

And a linear classifier (for example, perceptron) uses the classification rule:

$$\arg\max_y \sum_{i=0}^{n} w_{y,i} \cdot f_i \qquad \text{where } f_0 \text{ is a bias feature that is always 1 for all data}$$

(a) (8 pt) For a naive Bayes classifier with binary-valued features, i.e. $f_i \in \{0, 1\}$, prove that it is also a linear classifier by defining weights $w_{y,i}$ (for $i = 0, \ldots, n$) such that both decision rules above are equivalent. The weights should be expressed in terms of the naive Bayes probabilities — $P(y)$, $P(F_i = 1|y)$, and $P(F_i = 0|y)$. Assume that all these probabilities are non-zero.

Solution:

$$\begin{aligned}
\arg\max_y P(y) \prod_{i=1}^{n} P(f_i|y) &= \arg\max_y \ \log P(y) \ + \ \sum_{i=1}^{n} \log P(f_i|y) \\
&= \arg\max_y \ \log P(y) \ + \ \sum_{i=1}^{n} [f_i \log P(F_i = 1|y) + (1 - f_i) \log P(F_i = 0|y)] \\
&= \arg\max_y \ \log P(y) + \sum_{i=1}^{n} \log P(F_i = 0|y) \ + \ \sum_{i=1}^{n} f_i \log \frac{P(F_i = 1|y)}{P(F_i = 0|y)}
\end{aligned}$$

This is clearly equivalent to a linear classifier with the weights:

$$\begin{aligned}
w_{y,0} &= \log P(y) + \sum_{i=1}^{n} \log P(F_i = 0|y) \\
w_{y,i} &= \log \frac{P(F_i = 1|y)}{P(F_i = 0|y)} \qquad \text{For } i = 1, \ldots, n
\end{aligned}$$

(b) (2 pt) For the training set below with binary features F_1 and F_2 and label Y, name a smoothing method that would estimate a naive Bayes model that would correctly classify all four data, or answer "impossible" if there is no smoothing method that would give appropriate distributions $P(Y)$ and $P(F_i|Y)$. Briefly justify your answer.

F_1	F_2	Y
0	0	0
0	1	1
1	0	1
1	1	0

Solution:

The data is clearly not linearly separable. Since we know from the previous problem that a naive Bayes classifier on binary-valued features can be re-written as a linear classifier, no Naive Bayes classifier can exist for this data.

5. (17 points) Blind Connect Three

In Connect Three, players alternate dropping pieces into one of four columns. A player wins by having three consecutive pieces of their color either horizontally, vertically, or diagonally. Assume columns have infinite height. A dropped piece *always* occupies the lowest open space in that column.

You are playing the game blind-folded against a random opponent. You can't see the opponent's moves. However, you can *always* hear the opponent's piece sliding into place. When the opponent drops his piece along the edge of the board, it makes a *zing* sound; and when he drops it in one of the center two columns, it makes a *zang* sound. On the other hand, you know exactly the move which you have made. When a player wins, the referee stops the game and announces the winner.

We'll assume that you are representing your belief with a particle filter where each particle is a complete description of the board. Also, assume that you are the first player and that you are playing white. The only observations you get are the *zing* and the *zang* of the opponent pieces falling down.

(a) (8 pt) For each pair of particles below, answer *true* if it is possible to have both particles simultaneously present together in your particle filter immediately after a resampling step. If you answer *false*, *justify your answer for full credit.*

i. (*true* or *false*)

Solution:
False. There is no uncertainty about the number of moves played.

ii. (*true* or *false*)

Solution:
False. There is no uncertainty about the number of opponent pieces along the edge versus the center.

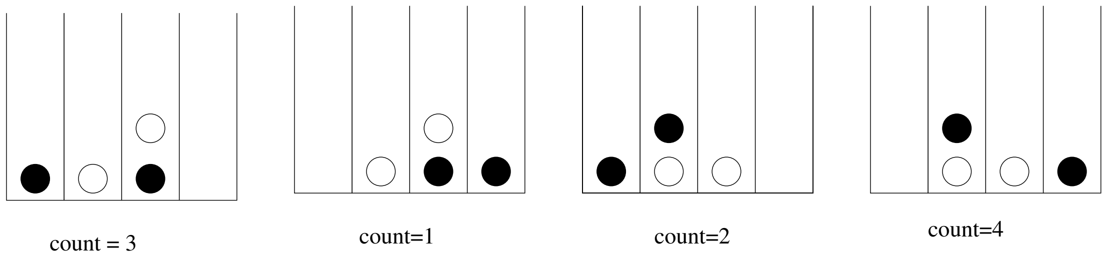
iii. (*true* or *false*)

Solution:
True.

iv. (*true* or *false*)

Solution:
False. There is no uncertainty about whether the game is over. (In the second particle, white has three in a row before black's last move)

Now, assume that you have the following 10 particles in your particle filter after two rounds of moves, where the count indicates a number of duplicate particles for the same game state. Your utility is 1 if you win on the next move, and 0 if you do not.
Note: even if you win eventually, your utility is 0 unless you win in one move.



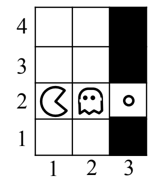
(b) (2 pt) According to your particle filter, what is your belief that it is possible to win in the next move?
Solution:
A total of $2 + 4 = 6$ particles out of 10 are winnable, so you believe that you can possibly win with probability 0.6.

(c) (3 pt) What is your maximum expected utility?
Solution:
The left and right columns are the only moves which have a non-zero probability of winning. Out of them, the left-most column has a greater probability of winning. MEU = .4.

(d) (4 pt) What is the value of peeking at the board before moving?
Solution:
The odds that you are in a winning position is $\frac{2+4}{10}$. Hence your expected MEU after peeking will be .6. The value of peeking is thus $.6 - .4 = .2$.

6. (22 points) The Last Dot

Pacman is one dot away from summer vacation. He just has to outsmart the ghost, starting in the game state to the right. Pacman fully observes every game state and will move first. The ghost only observes the starting state, the layout, its own positions after each move, and whether or not Pacman is in a middle row (2 and 3) or an edge row (1 and 4). The ghost probabilistically tracks what *row* pacman is in, and it moves around in column 2 to block Pacman.



Starting state

Details of the ghost agent:

- The ghost knows that Pacman starts in row 2. It *assumes* Pacman changes rows according to:
  - In rows 1 and 4, Pacman stays in the same row with probability $\frac{9}{10}$ and moves into the adjacent center row otherwise.
  - In rows 2 and 3, Pacman moves up with probability $\frac{1}{2}$, down with probability $\frac{3}{8}$, and stays in the same row with probability $\frac{1}{8}$.
- After each of Pacman's turns (but before the ghost moves), the ghost *observes* whether Pacman is in a middle row (2 and 3) or an edge row (1 and 4).
- Each turn, the ghost moves up, moves down or stops
- The ghost moves toward the row that most probably contains Pacman according to its model and observations. If the ghost believes Pacman is most likely in its current row, then it will stop.

(a) **(3 pt)** According to the ghost's *assumed* transition and emission models, what will be the ghost's belief distribution over Pacman's row if Pacman moves up twice, starting from the game state shown?

Row 1	Row 2	Row 3	Row 4

Solution:
The ghost believes Pacman is in row 4 with probability $\frac{16}{19}$, and in row 1 with probability $\frac{3}{19}$.

(b) **(3 pt)** If Pacman stops twice in a row (starting from the start state), **where** will the ghost be according to its policy? **With what probability** does the ghost believe it is in the same row as Pacman?

Ghost position	Ghost belief probability

Solution:
After one observation that Pacman is in the middle, the ghost's belief distribution will be $B(2) = \frac{1/8}{1/8+1/2} = \frac{1}{5}$ and $B(3) = \frac{1/2}{1/8+1/2} = \frac{4}{5}$. After a second observation, we have $B(2) = \frac{4/5*3/8+1/5*1/8}{4/5*3/8+1/5*1/8+4/5*1/8+1/5*1/2}$, which works out to $\frac{13}{21}$. Thus, the ghost will be in position (2,2) (row 2), because $\frac{13}{21} > \frac{1}{2}$.

(c) (2 pt) If Pacman moves down, then up repeatedly, alternating between positions (1,1) and (1,2), what will the ghost's belief distribution be about Pacman's row after n moves, as n goes to infinity?

Hint: you may want to consider multiple cases.

Solution:

If n is even, $P(R_n = (1,2)) = 1$. If n is odd, $P(R_n = (1,1)) = 1$.

(d) (6 pt) If Pacman moves up, then down repeatedly, alternating between positions (1,3) and (1,2), what will the ghost's belief distribution be about Pacman's row after n moves as n goes to infinity?

Hint: if $ax^2 + bx + c = 0$, then $x = \frac{-b \pm \sqrt{b^2-4ac}}{2a}$.

Solution:

Solving for $B(3)$ as x, we get $x = \frac{1/8*x+1/2*(1-x)}{1/8*x+1/2*(1-x)+3/8*x+1/8*(1-x)}$, which gives $4 - \sqrt{12} = 0.536$. Likewise, $B(2) = \sqrt{12} - 3 = 0.464$.

(e) (4 pt) Describe the state space for a *search problem* that Pacman could solve to find the shortest safe path to the food.

Solution:

The state includes the game state and the ghost's belief distribution. Note that without the ghost's belief distribution, you cannot tell where the ghost will go next based on the state, so you cannot correctly define the successor function.

(f) (4 pt) List the action sequence that Pacman should follow in order to reach the food in as few steps as possible. *Briefly* justify your answer.

Solution:

Pacman: stop, down, right, up, right. Ghost would go: up, up, stop, down.