You have 110 minutes to complete this exam. You are only allowed to use your textbook, your notes, your assignments and solutions to those assignments during this exam. If you find that you are spending a large amount of time on a difficult question, skip it and return to it when you’ve finished some of the easier questions. Total marks for this exam is 81.

<table>
<thead>
<tr>
<th>Section</th>
<th>Marks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-Midterm</td>
<td>/ 32</td>
</tr>
<tr>
<td>Probability</td>
<td>/ 10</td>
</tr>
<tr>
<td>Bayesian Networks</td>
<td>/ 39</td>
</tr>
<tr>
<td>Total</td>
<td>/ 81</td>
</tr>
</tbody>
</table>
Section I: Pre-Midterm questions [32 points]

1. Circle true or false below each question.
a) Every agent is rational in an unobservable environment. [2 points]

   True   False

b) An environment in which the agent must find a series of actions to achieve a goal is called dynamic. [2 points]

   True   False

c) A model-based reflex agent can outperform a simple reflex agent because it represents and maintains the state of the world. [2 points]

   True   False

d) If all step costs = 1, then breadth-first search and uniform cost search will find the same solution. [2 points]

   True   False

e) A* search with h(n)=0 for all nodes n is the same as uniform cost search. [2 points]

   True   False

f) If heuristic h1 dominates heuristic h2 and both heuristics are consistent, then we should choose to use h2. [2 points]

   True   False

g) The number of tiles out of place in the 8-tile-puzzle is an admissible heuristic because its value is 0 in the goal state. [2 points]

   True   False

h) Beam search with a beam width greater than 1 (i.e. k>1) is equivalent to simulated annealing with random restarts. [2 points]

   True   False

i) Gradient descent is only applicable to problems in which the objective function is differentiable. [2 points]

   True   False
j) The expectiminimax algorithm is only applicable to problems in which the game tree is too large to fit in memory. [2 points]

   True  False

k) The minimax algorithm is only applicable in deterministic environments. [2 points]

   True  False

l) The primary reason to use alpha-beta pruning is to increase the quality of the solutions found by the search algorithm. [2 points]

   True  False

m) The sentence \((A \land B) \lor (C \land D)\) is in conjunctive normal form. [2 points]

   True  False

n) The resolution algorithm is essentially a proof by contradiction. [2 points]

   True  False

o) The sentence \((A \lor B) \land (\neg A \lor \neg B)\) is satisfiable. [2 points]

   True  False

p) \(\neg(A \Rightarrow B) \equiv A \land \neg B\). [2 points]

   True  False
Section II: Probability [10 points]

2. You are conducting a political poll in Corvallis on the upcoming election. Having asked a large number of people what political party they belong to and whether they will vote for Donald Trump, you find that the probability of a person reporting that they will vote for Trump given that they are a Republican is 0.65, given that they are a Democrat is 0.05, and given that they are an Independent is 0.4. Based on the general population of Corvallis, the prior probability of being registered as a Democrat is 0.5, as an Independent is 0.35, and as a Republican is 0.15. Given that a particular respondent said they would NOT vote for Trump, what is the probability that they are registered as an Independent? [10 points]

Let T be a random variable for “voting for Trump” taking values true and false. Let A be a random variable for party affiliation, taking values Democrat (d), Republican (r), and Independent (i).

\[ P(T|A=r) = 0.65; \quad P(T|A=d) = 0.05; \quad P(T|A=i) = 0.4 \]
\[ P(A=d) = 0.5; \quad P(A=i) = 0.35; \quad P(A=r) = 0.15 \]

\[
P(A = i | \neg T) = \frac{P(\neg T | A = i)P(A = i)}{P(\neg T)} = \frac{(1 - P(T | A = i)) \times 0.35}{P(\neg T)} = \frac{(1 - 0.4) \times 0.35}{0.6 \times 0.35} = \frac{1 - \sum_A P(T, A)}{1 - \sum_A P(T, A)} = \frac{0.21}{1 - (P(T | A = r)P(A = r) + P(T | A = i)P(A = i) + P(T | A = d)P(A = d))} = \frac{0.21}{1 - (0.65 \times 0.15 + 0.4 \times 0.35 + 0.05 \times 0.5)} = \frac{0.21}{1 - 0.2625} = 0.285
\]
Section III: Bayesian Networks [39 points]

You are studying beavers at locations in the Willamette Valley. At each location in your study, you record whether the location is forested (F), whether it has a stream (S), and whether or not you detected a beaver at the location (D). Note that depending on how you collect your data, even if a beaver is present at the location (B), you might not detect it (i.e. B=true but D=false). You use the following Bayes net structure to model your data:

3. Your first task is to estimate the parameters of the conditional probability tables. For this, you use data on beavers with GPS collars (so you always know their true presence status B) and also visual surveys for D (in which you may not be able to see a beaver, e.g. if it’s underwater). A subset of the data you collect looks like this:

<table>
<thead>
<tr>
<th>Stream (S)</th>
<th>Forest (F)</th>
<th>Beaver Present (B)</th>
<th>Detected (D)</th>
</tr>
</thead>
<tbody>
<tr>
<td>True</td>
<td>True</td>
<td>True</td>
<td>False</td>
</tr>
<tr>
<td>True</td>
<td>True</td>
<td>True</td>
<td>True</td>
</tr>
<tr>
<td>True</td>
<td>False</td>
<td>True</td>
<td>True</td>
</tr>
<tr>
<td>False</td>
<td>True</td>
<td>False</td>
<td>False</td>
</tr>
<tr>
<td>True</td>
<td>False</td>
<td>False</td>
<td>False</td>
</tr>
<tr>
<td>True</td>
<td>True</td>
<td>True</td>
<td>False</td>
</tr>
<tr>
<td>False</td>
<td>False</td>
<td>False</td>
<td>False</td>
</tr>
</tbody>
</table>

a) Using this dataset, estimate the conditional probabilities in the table for S. The structure of the table is implied by the structure of the Bayesian network. Use Uniform Dirichlet priors as in Programming Assignment #3. [2 points]

<table>
<thead>
<tr>
<th>S</th>
<th>P(S)</th>
</tr>
</thead>
<tbody>
<tr>
<td>True</td>
<td>$\frac{5 + 1}{7 + 2} = \frac{6}{9} = \frac{2}{3}$</td>
</tr>
<tr>
<td>False</td>
<td>$\frac{2 + 1}{7 + 2} = \frac{3}{9} = \frac{1}{3}$</td>
</tr>
</tbody>
</table>

b) Using this dataset, estimate the conditional probabilities in the table for D. The structure of the table is implied by the structure of the Bayesian network. Use Uniform Dirichlet priors as in Programming Assignment #3. [4 points]

| B   | D   | P(D|B) |
|-----|-----|------|
| False| False| $\frac{3 + 1}{3 + 2} = \frac{4}{5}$ |
| False| True | $\frac{0 + 1}{3 + 2} = \frac{1}{5}$ |
| True | False| $\frac{2 + 1}{4 + 2} = \frac{3}{6} = \frac{1}{2}$ |
| True | True | $\frac{2 + 1}{4 + 2} = \frac{3}{6} = \frac{1}{2}$ |
4. Now, suppose you use the full dataset (not just the subset above) and get the following conditional probability tables (structure repeated for convenience):

<table>
<thead>
<tr>
<th>S</th>
<th>P(S)</th>
<th>F</th>
<th>P(F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>False</td>
<td>0.2</td>
<td>False</td>
<td>0.3</td>
</tr>
<tr>
<td>True</td>
<td>0.8</td>
<td>True</td>
<td>0.7</td>
</tr>
</tbody>
</table>

| S  | F  | B    | P(B|S,F) |
|----|----|------|---------|
| False | False | False | 0.95    |
| False | False | True  | 0.05    |
| False | True  | False | 0.9     |
| False | True  | True  | 0.1     |
| True  | False | False | 0.55    |
| True  | False | True  | 0.45    |
| True  | True  | False | 0.25    |
| True  | True  | True  | 0.75    |

| B  | D  | P(D|B) |
|----|----|-------|
| False | False | 0.99  |
| False | True  | 0.01  |
| True  | False | 0.4   |
| True  | True  | 0.6   |

a) Using these CPTs and the Bayes net structure, compute the probability that a beaver is present at a location in the absence of any other information: \( P(B=\text{true}) \). Show your work. [10 points]

\[
P(B = \text{true}) = \sum_{S,F,D} P(S,F,B,D)
\]

\[
= \sum_{S,F,D} P(S)P(F)P(B = \text{true}|S,F)P(D|B = \text{true})
\]

\[
= \sum_{S} P(S) \sum_{F} P(F)P(B = \text{true}|S,F) \sum_{D} P(D|B = \text{true})
\]

\[
= \sum_{S} P(S) \sum_{F} P(F)P(B = \text{true}|S,F)
\]

\[
= P(S = \text{true})(P(F = \text{true})P(B = \text{true}|S = \text{true},F = \text{true}) + P(F = \text{false})P(B = \text{true}|S = \text{true},F = \text{false})))
\]

\[
+ P(S = \text{false})(P(F = \text{true})P(B = \text{true}|S = \text{false},F = \text{true}) + P(F = \text{false})P(B = \text{true}|S = \text{false},F = \text{false}))
\]

\[
= 0.8(0.7 \times 0.75 + 0.3 \times 0.45) + 0.2(0.7 \times 0.1 + 0.3 \times 0.05)
\]

\[
= 0.545
\]
b) Suppose you visit a forested location with a stream and do not detect a beaver. Given this information, would you predict that a beaver is present at the location or not? Show your work. (Structure and tables repeated for convenience.) [10 points]

\[
P(B = \text{true}|S = \text{true}, F = \text{true}, D = \text{false}) \quad \text{or} \quad P(B = \text{false}|S = \text{true}, F = \text{true}, D = \text{false})?
\]

\[
P(B = \text{true}|S = \text{true}, F = \text{true}, D = \text{false}) = a \times P(S = \text{true})P(F = \text{true})P(B = \text{true}|S = \text{true}, F = \text{true}) \times P(D = \text{false}|B = \text{true})
\]

\[
= a \times 0.8 \times 0.7 \times 0.75 \times 0.4
\]

\[
= a \times 0.168
\]

\[
P(B = \text{false}|S = \text{true}, F = \text{true}, D = \text{false}) = a \times P(S = \text{true})P(F = \text{true})P(B = \text{false}|S = \text{true}, F = \text{true}) \times P(D = \text{false}|B = \text{false})
\]

\[
= a \times 0.8 \times 0.7 \times 0.25 \times 0.99
\]

\[
= a \times 0.1386
\]

Therefore, predict $B=\text{true}$.
5. A new study suggests that on cloudy days, you are much less likely to detect a beaver, even if it is present. Draw a new Bayesian network structure that adds the node C (for cloudy) to the existing four nodes to represent this information. Your structure should have I(B,C)=true, I(D,C)=false, and I(B,C|D)=false. C should also be independent of S and F, and the rest of the independence relationships for the original four nodes should remain the same. [4 points]
6. Use the Bayesian network below to determine whether or not the following conditional independence relationships hold or not. Show the blocked/unblocked paths for partial credit.

![Bayesian network diagram]

a) $I(A,E|B)$ [3 points]

A-B-E: blocked by B
A-D-G-E: blocked by G
A-B-F-H-E: blocked by B and H

True.

b) $I(A,C|\{D,E,F\})$ [3 points]

A-B-F-C: not blocked
A-B-E-H-F-C: blocked by H and F
A-D-G-E-H-F-C: blocked by D, G, E, H, F
A-D-G-E-B-F-C: blocked by D, G, E

False.

c) $I(C,G|\{E,F\})$ [3 points]

C-F-H-E-G: blocked by F, E, H
C-F-B-E-G: blocked by E
C-F-B-A-D-G: not blocked

False.