Q1. Bayes Nets and Joint Distributions

(a) Write down the joint probability distribution associated with the following Bayes Net. Express the answer as a product of terms representing individual conditional probabilities tables associated with this Bayes Net:

(b) Draw the Bayes net associated with the following joint distribution:
\[ P(A) \cdot P(B) \cdot P(C|A, B) \cdot P(D|C) \cdot P(E|B, C) \]

(c) Do the following products of factors correspond to a valid joint distribution over the variables A, B, C, D? (Circle FALSE or TRUE.)

(i) FALSE TRUE \[ P(A) \cdot P(B) \cdot P(C|A) \cdot P(C|B) \cdot P(D|C) \]

(ii) FALSE TRUE \[ P(A) \cdot P(B|A) \cdot P(C) \cdot P(D|B, C) \]

(iii) FALSE TRUE \[ P(A) \cdot P(B|A) \cdot P(C) \cdot P(C|A) \cdot P(D) \]

(iv) FALSE TRUE \[ P(A|B) \cdot P(B|C) \cdot P(C|D) \cdot P(D|A) \]
(d) What factor can be multiplied with the following factors to form a valid joint distribution? (Write “none” if the given set of factors can’t be turned into a joint by the inclusion of exactly one more factor.)

(i) \( P(A) \cdot P(B|A) \cdot P(C|A) \cdot P(E|B, C, D) \)

(ii) \( P(D) \cdot P(B) \cdot P(C|D, B) \cdot P(E|C, D, A) \)

(e) Answer the next questions based off of the Bayes Net below:

All variables have domains of \([-1, 0, 1]\)

![Bayes Net Diagram]

(i) Before eliminating any variables or including any evidence, how many entries does the factor at G have?

(ii) Now we observe \( e = 1 \) and want to query \( P(D|e = 1) \), and you get to pick the first variable to be eliminated.

- Which choice would create the largest factor \( f_i \)?

- Which choice would create the smallest factor \( f_i \)?
Q2. Bayes Nets: Sampling

Consider the following Bayes Net, where we have observed that $B = +b$ and $D = +d$.

(a) Consider doing Gibbs sampling for this example. Assume that we have initialized all variables to the values $+a, +b, +c, +d$. We then unassign the variable $C$, such that we have $A = +a, B = +b, C = ?, D = +d$. Calculate the probabilities for new values of $C$ at this stage of the Gibbs sampling procedure.

$$P(C = +c \text{ at the next step of Gibbs sampling}) =$$

$$P(C = -c \text{ at the next step of Gibbs sampling}) =$$

(b) Consider a sampling scheme that is a hybrid of rejection sampling and likelihood-weighted sampling. Under this scheme, we first perform rejection sampling for the variables $A$ and $B$. We then take the sampled values for $A$ and $B$ and extend the sample to include values for variables $C$ and $D$, using likelihood-weighted sampling.

(i) Below is a list of candidate samples. Mark the samples that would be rejected by the rejection sampling portion of the hybrid scheme.

- $-a -b$
- $+a +b$
- $+a -b$
- $-a +b$

(ii) To decouple from part (i), you now receive a new set of samples shown below. Fill in the weights for these samples under our hybrid scheme.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>$-a +b -c +d$</td>
<td></td>
</tr>
<tr>
<td>$+a +b -c +d$</td>
<td></td>
</tr>
<tr>
<td>$+a +b -c +d$</td>
<td></td>
</tr>
<tr>
<td>$-a +b +c +d$</td>
<td></td>
</tr>
<tr>
<td>$+a +b +c +d$</td>
<td></td>
</tr>
</tbody>
</table>

(iii) Use the weighted samples from part (ii) to calculate an estimate for $P(+a \mid b, +d)$. The estimate of $P(+a \mid b, +d)$ is
(e) We now attempt to design an alternative hybrid sampling scheme that combines elements of likelihood-weighted and rejection sampling. For each proposed scheme, indicate whether it is valid, i.e. whether the weighted samples it produces correctly approximate the distribution \( P(A, C | b, +d) \).

(i) **First collect a likelihood-weighted sample for the variables A and B. Then switch to rejection sampling for the variables C and D. In case of rejection, the values of A and B and the sample weight are thrown away. Sampling then restarts from node A.**

- Valid
- Invalid

(ii) **First collect a likelihood-weighted sample for the variables A and B. Then switch to rejection sampling for the variables C and D. In case of rejection, the values of A and B and the sample weight are retained. Sampling then restarts from node C.**

- Valid
- Invalid