Dong Hu

HOMEWORK 1

1 CSP

Problem 1

Consider the problem of placing k knights on an $n \times n$ chessboard such that no two knights are attacking each other, where k is given and $k \leq n^2$

a. Choose a CSP formulation. In your formulation, what are the variables?

   Answer: The variables are the positions on the $n \times n$ chessboard.

b. What are the possible values of each variable?

   Answer: The possible values of each variable are either True of False. The value is True if the position is occupied my a knight, False otherwise.

c. What sets of variables are constrained, and how?

   Answer: Every pair of squares separated by a knights move is constrained, such that both cannot be occupied. Furthermore, the entire set of squares is constrained, such that the total number of occupied squares should be k.

.

2 Probability

Problem 2

Prove the chain rule. That is, for any probabilistic model composed of random variables $X_1, \ldots, X_n$ and any values $x_1, \ldots, x_n$, we have:

$$p(x_1, \ldots, x_n) = \prod_{i=1}^{n} p\left(x_i | x_1, \ldots, x_{i-1}\right)$$

We use prove by induction to complete the proof.

1. Base Case: $n = 2$

$$p(x_1, x_2) = p(x_2|x_1)p(x_1) = \prod_{i=1}^{2} p\left(x_i | x_1, \ldots, x_{i-1}\right)$$

2. Induction Step:

   We first suppose the statement is true for $p(x_1, \ldots, x_k) = \prod_{i=1}^{k} p\left(x_i | x_1, \ldots, x_{i-1}\right)$, then we want to prove that it still holds true for $p(x_1, \ldots, x_{k+1}) = \prod_{i=1}^{k+1} p\left(x_i | x_1, \ldots, x_{i-1}\right)$, where $2 \leq k < n$.

   Suppose $p(x_1, \ldots, x_k) = p(a) = \prod_{i=1}^{k} p\left(x_i | x_1, \ldots, x_{i-1}\right)$, by the defination of product rule,

$$\begin{aligned} p(x_1, \ldots, x_{k+1}) &= p(a, x_{k+1}) = p(x_{k+1}|a)p(a) \\ &= p(x_{k+1}|x_1, \ldots, x_k) \prod_{i=1}^{k} p\left(x_i | x_1, \ldots, x_{i-1}\right) \\ &= \prod_{i=1}^{k+1} p\left(x_i | x_1, \ldots, x_{i-1}\right) \end{aligned}$$

3. Conclusion:

   Since we successfully prove the base case, and prove that $p(x_1, \ldots, x_k) = \prod_{i=1}^{k} p\left(x_i | x_1, \ldots, x_{i-1}\right)$ implies $p(x_1, \ldots, x_{k+1}) = \prod_{i=1}^{k+1} p\left(x_i | x_1, \ldots, x_{i-1}\right)$. ■

Problem 3

Prove that the two definitions of conditional independence of random variables are equivalent. Let X, Y, Z be random variables. The two definitions are:

- Definition 1: X and Y are conditionally independent given Z if for any value x of X,any value y of Y, and any value z of Z, the following holds: $p(x, y|z) = p(x|z) \times p(y|z)$.
- Definition 2: X and Y are conditionally independent given Z if for any value x of X, any value y of Y, and any value z of Z, the following holds: $p(x|y, z) = p(x|z)$.

$$p(x, y|z) = \frac{p(x, y, z)}{p(z)} = \frac{p(x|y, z) \times p(y, z)}{p(z)} = p(x|y, z) \times \frac{p(y, z)}{p(z)} = p(x|y, z) \times p(y|z)$$

Apply definition 2 on the above equation:

$$p(x|y, z) = p(x|y, z) \times p(y|z) = p(x|z) \times p(y|z)$$

The result consists with definition 1, Hence they are equivalent. ■

Problem 4

Let X, Y, Z be random variables. Prove or disprove the following statements. (That means, you need to either write down a formal proof, or give a counterexample.)

- Statement 1: If X and Y are (unconditionally) independent, is it true that X and Y are conditionally independent given Z?

  Answer: False.

  $X \rightarrow Z \leftarrow Y$, where X and Y are unconditionally independent, wheares they are dependent after Z is given.

- Statement 2: If X and Y are conditionally independent given Z, is it true that X and Y are (unconditionally) independent?

  Answer: False.

  $X \leftarrow Z \rightarrow Y$, where X and Y are conditionally independent when Z is given, however they are unconditionally dependent.

Bayesian networks

Problem 5

1. Conditional dependence. Active path: $T \leftarrow I \rightarrow U$
2. Conditional dependence. Active path: $T \rightarrow \underline{E} \leftarrow U$
3. Conditional independence. No active path between T and U
4. Conditional dependence. Active path: $E \leftarrow T \leftarrow I \rightarrow \underline{U} \leftarrow H$
5. Conditional independence. No active path between E and H
6. Conditional dependence. Active path: $\begin{array}{c} I \rightarrow U \leftarrow H \\ \downarrow \\ \underline{E} \end{array}$
7. Conditional independence. No active path between I and H
8. Conditional independence. No active path between T and H
9. Conditional dependence. Active path: $T \rightarrow \underline{E} \leftarrow U \leftarrow H$
10. Conditional dependence. Active path: $T \leftarrow I \rightarrow \underline{U} \leftarrow H$

Problem 6

Since

$$P(+u|+e) = \frac{P(+u,+e)}{p(+e)}$$

So we Calculate $P(+u|+e)$ first

$$\begin{aligned} P(+u|+e) &= \sum_{i,t,h}(+u,+e,i,t,h) \\ &= \sum_{i,t,h} P(i)\cdot P(t|i)\cdot P(+u|i,h)\cdot P(h)\cdot P(+e|+u,t) \\ &= \sum_{h} P(h)\sum_{i} P(i)\cdot P(+u|i,h)\cdot \sum_{t} P(t|i)\cdot P(+e|+u,t) \end{aligned} \tag{1}$$

So if we take

$$f_2(h,i) = \sum_t P(t|i)\cdot P(+e|+u,t)$$

$$\begin{aligned} f_2(h,+i) &= P(+t|+i)\cdot P(+e|+u,+t) + P(-t|+i)\cdot P(+e|+u,-t) \\ &= 0.8\cdot 0.9 + 0.2\cdot 0.7 \\ &= 0.86 \end{aligned} \tag{2}$$

$$\begin{aligned} f_2(h,-i) &= P(+t|-i)\cdot P(+e|+u,+t) + P(-t|-i)\cdot P(+e|+u,-t) \\ &= 0.5\cdot 0.9 + 0.5\cdot 0.7 \\ &= 0.80 \end{aligned} \tag{3}$$

Also

$$f_1(h) = \sum_i P(i)\cdot P(+u|i,h)\cdot f_2(h,i)$$

$$\begin{aligned} f_1(+h) &= P(+i)\cdot P(+u|+i,+h)\cdot f_2(+h,+i) + P(-i)\cdot P(+u|-i,+h)\cdot f_2(+h,-i) \\ &= 0.7\cdot 0.9\cdot 0.86 + 0.3\cdot 0.5\cdot 0.80 \\ &= 0.6618 \end{aligned} \tag{4}$$

$$\begin{aligned} f_1(-h) &= P(+i)\cdot P(+u|+i,-h)\cdot f_2(-h,+i) + P(-i)\cdot P(+u|-i,-h)\cdot f_2(-h,-i) \\ &= 0.7\cdot 0.3\cdot 0.86 + 0.3\cdot 0.1\cdot 0.80 \\ &= 0.2046 \end{aligned} \tag{5}$$

Then

$$\begin{aligned} P(+u,+e) &= \sum_h P(h)\cdot f_1(h) \\ &= p(+h)\cdot f_1(+h) + p(-h)\cdot f_1(-h) \\ &= 0.6\cdot 0.6618 + 0.4\cdot 0.2046 \\ &= 0.47892 \end{aligned} \tag{6}$$

Then we Calculate $P(-u|+e)$

$$
\begin{aligned}
P(-u|+e) &= \sum_{i,t,h}(-u,+e,i,t,h) \\
&= \sum_{i,t,h} P(i)\cdot P(t|i)\cdot P(-u|i,h)\cdot P(h)\cdot P(+e|-u,t) \\
&= \sum_{h} P(h)\sum_{i} P(i)\cdot P(-u|i,h)\cdot \sum_{t} P(t|i)\cdot P(+e|-u,t)
\end{aligned} \tag{7}
$$

So if we take

$$f_2(h,i) = \sum_t P(t|i)\cdot P(+e|-u,t)$$

$$
\begin{aligned}
f_2(h,+i) &= P(+t|+i)\cdot P(+e|-u,+t) + P(-t|+i)\cdot P(+e|-u,-t) \\
&= 0.8\cdot 0.5 + 0.2\cdot 0.3 \\
&= 0.46
\end{aligned} \tag{8}
$$

$$
\begin{aligned}
f_2(h,-i) &= P(+t|-i)\cdot P(+e|-u,+t) + P(-t|-i)\cdot P(+e|-u,-t) \\
&= 0.5\cdot 0.5 + 0.5\cdot 0.3 \\
&= 0.4
\end{aligned} \tag{9}
$$

Also

$$f_1(h) = \sum_i P(i)\cdot P(-u|i,h)\cdot f_2(h,i)$$

$$
\begin{aligned}
f_1(+h) &= P(+i)\cdot P(-u|+i,+h)\cdot f_2(+h,+i) + P(-i)\cdot P(-u|-i,+h)\cdot f_2(+h,-i) \\
&= 0.7\cdot 0.1\cdot 0.46 + 0.3\cdot 0.5\cdot 0.40 \\
&= 0.0922
\end{aligned} \tag{10}
$$

$$
\begin{aligned}
f_1(-h) &= P(+i)\cdot P(-u|+i,-h)\cdot f_2(-h,+i) + P(-i)\cdot P(-u|-i,-h)\cdot f_2(-h,-i) \\
&= 0.7\cdot 0.7\cdot 0.46 + 0.3\cdot 0.9\cdot 0.40 \\
&= 0.3334
\end{aligned} \tag{11}
$$

Then

$$
\begin{aligned}
P(-u,+e) &= \sum_h P(h)\cdot f_1(h) \\
&= P(+h)\cdot f_1(+h) + p(-h)\cdot f_1(-h) \\
&= 0.6\cdot 0.0922 + 0.4\cdot 0.3334 \\
&= 0.18868
\end{aligned} \tag{12}
$$

Then

$$
\begin{aligned}
P(+e) &= P(+u)\cdot P(+e|+u) + P(-u)\cdot P(+e|-u) \\
&= P(+u,+e) + P(-u,+e) \\
&= 0.18868 + 0.47892 \\
&= 0.6676
\end{aligned} \tag{13}
$$

Hence

$$P(+u|+e) = \frac{P(+u,+e)}{P(+e)} = 0.7112$$