1 CSP

Problem 1)

6.2)

a) The variables for this formulation are the k knights on the board.
b) The values for each variable can be the possible squares that are reachable by the night from its current position.
c) The variable constraints are that there cannot be intersecting sets of reachable squares for each variable, ensuring that no knights are able to attack in one move, and also that no two knights are in the same location.

---

2 Probability

Problem 2 (10 points.) Prove the chain rule. That is, for any probabilistic model composed of random variables $X_1, \dots, X_n$ and any values$x_1, \dots, x_n$, we have:

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

Problem 3 (10 points.) 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: $p(x|y, z) = p(x|z)$.

Problem 4 (bonus question 20 points.) 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?

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

Problem 2)

We can prove the chain rule by induction:

Base Case: $p(x_i) = \prod_{i=1}^{1} p(x_i) =: p(x_i)$

Inductive Step: We assume that for $n = k$; $p(x_1 \ldots x_k) = \prod_{i=1}^{k} p(x_i \mid x_1 \ldots x_{i-1})$ is true.

For $n = k + 1$,
$$= \prod_{i=1}^{k+1} p(x_i \mid x_{i-1} \ldots x_1)$$
$$= \left(\frac{p(x_1 \ldots x_{k+1})}{p(x_1 \ldots x_k)}\right) * p(x_1 \ldots x_k) = p(x_1 \ldots x_{k+1})$$

Therefore the case stands for $k + 1$.

Problem 3)

From $p(x, y \mid z) = p(x \mid z) * p(y \mid z)$ we can derive:

$$\frac{p(x,y,z)}{p(z)} = \frac{\frac{p(x,z)}{p(z)} * p(y,z)}{p(z)}$$
$$\frac{p(x,y,z)}{p(y,z)} = \frac{\frac{p(x,y)}{p(z)} * p(z)}{p(z)}$$
$$p(x \mid y, z) = p(x \vee z)$$

which shows equivalence to the second definition.

Problem 4)

Statement 1: The statement is false given the counterexample that if we flip two coins we can assign X to be the event that the first coin lands on heads, Y to be the event that the second coin lands on heads, and Z to be the event that the two coins landed on the same side. While X and Y are independent, given Z they are now conditionally dependent, so NO.

Statement 2: The statement is false given a counter example where X is the event that you go to the store to get groceries, and Y is the event that you know how to ride a bike. These events are independent until we introduce variable Z which is the event that the store is too far to walk and too short to drive, assuming bikes are the only remaining mode of transportation. Now X and Y given Z are dependent events which means they are independent without Z and dependent given Z, so NO.

3 Bayesian networks

We are going to take the perspective of an instructor who wants to determine whether a student has understood the material, based on the exam score. Figure 1 gives a Bayesian network for this. As you can see, whether the student

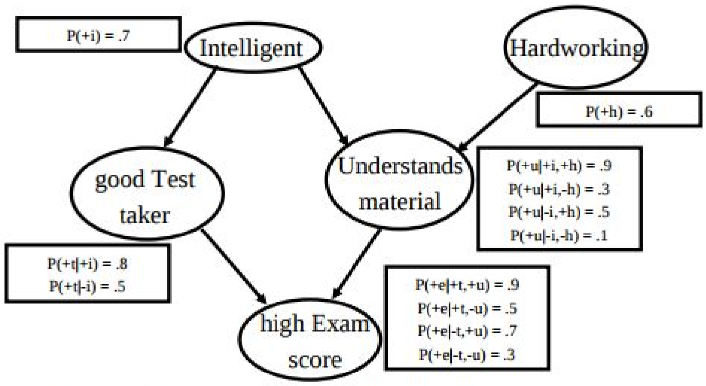

Figure 1: A Bayesian network representing what influences an exam score.

scores high on the exam is influenced both by whether she is a good test taker, and whether she understood the material. Both of those, in turn, are influenced by whether she is intelligent; whether she understood the material is also influenced by whether she is a hard worker.

Problem 5 (40 points.)

For the above Bayesian network, label the following statements about conditional independence as true or false. For this question, you should consider only the structure of the Bayesian network, not the specific probabilities. Explain each of your answers. Specifically, **when you claim conditional dependence, you must show an active path**.

1. T and U are independent.
2. T and U are conditionally independent given I, E, and H.
3. T and U are conditionally independent given I and H.
4. E and H are conditionally independent given U.
5. E and H are conditionally independent given U, I, and T.
6. I and H are conditionally independent given E.
7. I and H are conditionally independent given T.
8. T and H are independent.
9. T and H are conditionally independent given E.
10. T and H are conditionally independent given E and U.

Problem 5)

1. False, T and U are influenced by I.
2. False, T and U influence E making an active path.
3. True, T and U are conditionally independent given I and H.
4. False, E is influenced by $I \rightarrow T \rightarrow E$.
5. True, E and H are conditionally independent given I, T and U.

6. False, Given E the active path can be influenced by either I or H.
7. True, I and H are independent and are conditionally independent because H has no influence on T.
8. True, T and H are independent.
9. False, E is influenced by both T and H creating active paths $T \to E$ as well as $H \to U \to E$.
10. False, both T and U are influenced by I.

Problem 6)

Using variable elimination (by hand!), compute the probability that a student who did well on the test actually understood the material, that is, compute $P(+u|+e)$.

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

We are given $P(I),\ P(T|I),\ P(H),\ P(+u|I,H),\ P(+e|T,+u)$

and need to eliminate I, T, H

We eliminate H to get f_1

$f_1(I,+u) = f_1(+i,\ +u) + f_1(-i,\ +u)$

$= [P(+h)P(+u|+i,+h) + P(-h)P(+u|+i,-h)] + [P(+h)P(+u|-i,+h) + P(-h)P(+u|-i,-h)]$

$= (0.6 * 0.9 + 0.4 * 0.3) + (0.6 * 0.5 + 0.4 * 0.1)$

$= 0.66 + 0.34$

$= 1$

leaving $P(I),\ P(T|I),\ P(+e|T,+u),\ f_1(I,+u)$

We eliminate I to get f_2

$f_2(T,+u) = f_2(+t,+u) + f_2(-t,+u)$

$= [P(+i)P(+t|+i,)f_1(+i,\ +u) + P(-i)P(+t|-i,)f1(-i,\ +u)] + [P(+i)P(-t|+i,)f_1(+i,\ +u)$
$+ P(-i)P(-t|-i,)f_1(-i,\ +u)]$

$= [0.7 * 0.8 * 0.66 + 0.3 * 0.5 * 0.34] + [0.7 * 0.2 * 0.66 + 0.3 * 0.5 * 0.34]$

$= 0.4206 + 0.1434$

$= 0.564$

leaving $P(+e|T,+u),\ f_2(T,+u)$

We eliminate T to obtain f_3

$P(+e) = f_3(+e,+u) + f_3(+e,-u)$

$= [P(+e|+t,+u) * f_2(+t,+u) + P(+e|-t,+u),\ f_2(-t,+u)] + [P(+e|+t,-u) * f_2(-t,-u)$
$+ P(+e|-t,-u),\ f_2(-t,-u)]$

$= [0.9 * 0.4206 + 0.7 * 0.1434] + [0.5 * 0.4206 + 0.3 * 0.1434]$

$= 0.47892 + 0.18868$

$= 0.6676$

$$P(+u \mid +e) = \frac{P(+u,+e)}{P(+e)} = \frac{0.47892}{0.6676} = 0.71738$$