CSCI 4150: Introduction to Artificial Intelligence

Homework 1: Probability and Bayesian networks
Total points: 90. Bonus points: 10.

We only accept electronic submission at Submitty. Please try to ask questions on Piazza. If Piazza is not helpful, please contact the TAs.

1 Probability

Problem 1 (10 points.) 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:

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

Problem 2 (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 3 (bonus question 10 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 (5pts). If X and Y are (unconditionally) independent, is it true that X and Y are conditionally independent given Z?

Statement 2 (5pts). If X and Y are conditionally independent given Z, is it true that X and Y are (unconditionally) independent?

2 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.

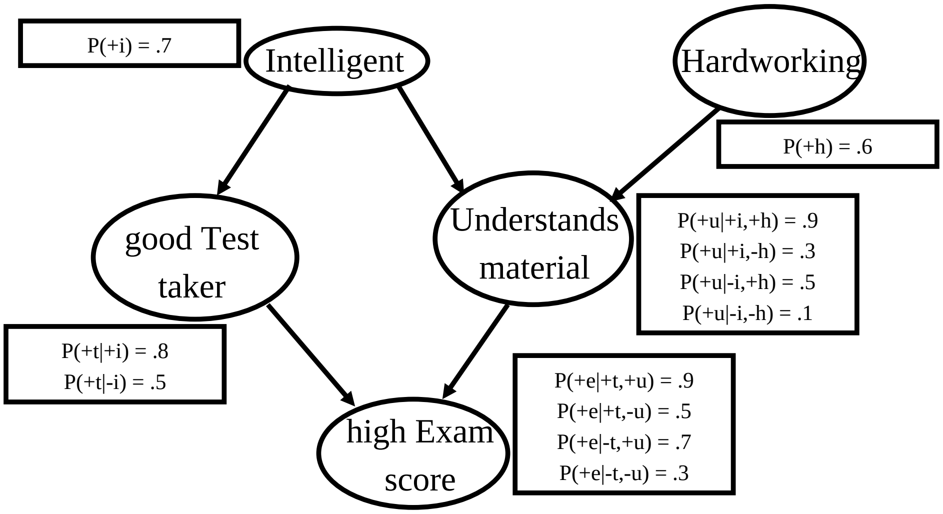

Figure 1: A Bayesian network representing what influences an exam score.

As you can see, whether the student 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 4 (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. **Please show all four steps and the reasoning after the four steps taught in the class, otherwise no partial score will be given.**

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 (30 points).

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