CPSC 322, Practice Exercise
Solutions to Variable Elimination

1 Directed Questions

- What is a factor? **Answer:** A factor is a representation of a function from a tuple of random variables into a number. A factor f on variables $X_1, ...X_j$ is written as $f(X_1, ..., X_j)$. $f(X_1 = v_1, X_2 = v_2, ..., X_J = v_j)$ is a number that is the value of f when each X_i has value v_i.
- What are the operations applied to factors in the variable elimination algorithm? **Answer:** Summing, multiplying, and assigning.
- What do we mean by the *belief network inference problem*?

  Answer: This is the problem of calculating the posterior distribution of some variable given some evidence.
- Define *elimination ordering*. **Answer:** This refers to the order in which variables are summed out.
- What are the three main steps of variable elimination?

  Answer:

  - Construct a factor for each conditional probability distribution
  - Eliminate each of the non-query variables:
    - if the variable is observed its value is set to the observed value in each of the factors in which the variable appears,
    - otherwise, the variable is summed out
  - Multiply the remaining factors and normalize

2 Belief Networks and Variable Elimination

Bill has noticed that his morning newspaper delivery has been sporadic. There are several relevant variables relating to whether or not the paper is delivered. Delivery is dependent on the paper having been successfully printed the previous night. Possible explanations for a paper not having been printed are a malfunction at the printing press, or the end of civilization as we know it.

Before continuing, write down the relevant variables for this scenario.

Answer: There are four variables, along the lines of $paperDelivered$, $paperPrinted$, $printerMalfunction$ and $endOfCivilization$.

Let's assign some probabilities. The prior probability of a printer malfunction is 0.05. Bill has been noticing some ominous signs of the apocalypse and so expects the end of civilization with a relatively high probability of 0.001. If the end of civilization is here, then the paper not be printed for sure. If there is a printing malfunction and no end of civilization, there is a probability

of 0.05 that the paper will be printed (this is non-zero because the malfunction might be fixed in time). If there is no malfunction and no end of civilization, there is a probability of 0.99 that the paper will be printed. If the paper is not printed it will not be delivered. If it is printed, there is a probability of 0.9 that it will be delivered. The fact that this probability is not 1 suggests that there are other possible causes for the paper not being delivered that we should eventually add to our belief network (e.g. the paperboy being sick).

- Construct this belief network in AISpace. Build the truth tables according to the probabilities above. **Answer:** Figure **??** illustrates the belief network for this example. The xml file `newspaper.xml` gives the AISpace representation.

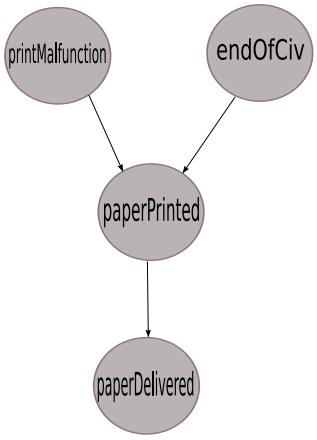

Figure 1: Belief Network For Newspaper Example

- Bill's paper fails to arrive one morning. He'd like to know the probability that civilization has ended. What are the observed variables, and what is the query variable? **Answer:** The observed variable is $paperDelivered$ and the query is $endOfCivilization$.
- What are the initial factors (prior to variable elimination)? **Answer:** $f0(printingMalfunction)$, $f1(endOfCivilization)$, $f2(printingMalfunction, endOfCivilization, paperPrinted)$, $f3(paperPrinted, paper1$
- Carry out variable elimination for this inference problem. Specify the elimination ordering you are using. Show each step of your work.

  Answer: The initial factors are given above. We can first do *assignment*. We assign the value $false$ to $paperDelivered$. $f3$ is removed and a new factor $f4(paperPrinted)$ is added. We can next sum out $printingMalfunction$, which removes factors $f0$ and $f2$. A new factor $f5(endOfCivilization, paperPrinted)$ is added. If we next sum out $paperPrinted$, then $f4$ and $f5$ are removed and $f6(endOfCivilization)$ is added. At this point we are left with just two factors, $f1(endOfCivilization)$ and $f6(endOfCivilization)$. Note that $f1$ corresponds to the prior for that variable. We then just need to multiply those final factors $f1$ and $f6$ to get $f7$ and normalize to get $f8$.
- What is the probability that civilization has ended given the observations? **Answer:** The probability is 0.00657. After multiplying the final factors, $f7$ give the value for $endOfCivilization = true$ as 0.001 and the value for $endOfCivilization = false$ as 0.15115. To normalize, we divide 0.001 by (0.001+0.15115) to get 0.00657.

- We were assuming that the query variable was not observed. If it *is* observed, does this simplify the inference problem? **Answer:** If the query variable is observed, we don't need to do any inference. The probability of the observed value is 1 and the probability of all other values is 0.

3 Learning Goals

You can:

- Build a belief network for a simple domain.
- Define factors. Derive new factors from existing factors.
- Carry out variable elimination by using factor representation and using the factor operations.
- Use techniques to simplify variable elimination.