CS 486 — Lecture 13: Variable Elimination Algorithms

1 VE Motivation

- For something like $P(B = b|W = t \wedge G = t), b \in t, f$, our shorthand is $P(B|w \wedge g)$.
- This is equal to:

$$P(B|w \wedge g) = \frac{P(B \wedge w \wedge g)}{P(b \wedge w \wedge g) + P(\neg b \wedge w \wedge g)}$$

- So for $P(B \wedge w \wedge g)$, we can say it's equal to:

$$\sum_a \sum_e \sum_r P(B)P(e)P(r|e)P(a|B \wedge e)P(w|a)P(g|a)$$

  Such an expression would take 47 operations to do!

- Note we can simplify this:

$$P(B) \sum_a P(w|a)P(g|a) \sum_e P(e)P(a|B \wedge e)$$

  Note the $P(r|e)$ term is removed as summing over r means it must equal 1.

- This would take *12* operations.

2 VE Algorithm

- Use dynamic programming (do calculations once if possible) and exploit conditional independence to reduce the number of operations required.
- VEA relies on factors and operations on factors.
- A factor is a function from some random variables to a number.
- For example, $f(X_1, X_2)$ could be $P(X_1 \wedge X_2)$ or $P(X_1|X_2)$.
- We define a factor for every conditional probability distribution in the Bayes net.
- We can *restrict* a factor by assigning a value to the variable in the factor.
- If one restricts until all values have assigned values in the factor, it is now just a number (obviously).
- *Sum out* operations sum out a variable:

$$(\sum_{X_1} f)(X_2, \cdots, X_j) = f(X_1 = v_1, \ldots, X_j) + \cdots + f(X_1 = v_k, \ldots, X_j)$$

- We can also *multiply* two factors together. The product of two factors $f_1(X, Y)$ and $f_2(Y, Z)$ where Y is in common will give $(f_1 \times f_2)(X, Y, Z)$.
- *Normalizing* divides each value by the sum of all the values. This ensures that we have a valid probability distribution (ie: if we had values of 0.2 and 0.6, we would want 0.25 and 0.75 as probabilities).
- The general VE algorithm:
  1. Construct a factor for each conditional probability.

2. Restrict observed variables to their observed values.
3. Eliminate each hidden variable via multiplying and summing out.
4. Multiply remaining factors.
5. Normalize the resulting factor.