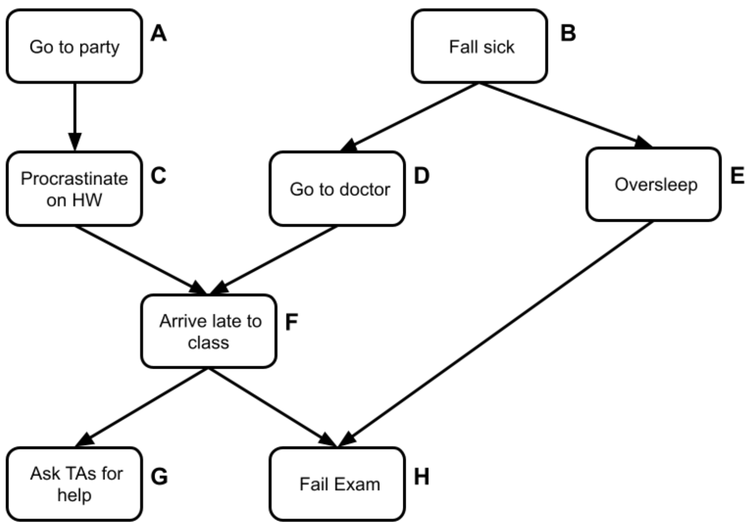
Learning Objectives

- To practice variable elimination on a sample Bayes net

Q1. Bayes Nets Inference

Suppose you are concerned about your social life, physical health, and procrastination and how it impacts your performance in class. You draw the following Bayes Net. The letters are provided for ease of reference to the variables.



(a) Given what we learned about inference, simplify the probability P(A,B,C,D) using the CPTs by pushing in the summations and summing to 1 when possible.

P(A,B,C,D) =

$\sum_e \sum_f \sum_g \sum_h P(A)P(B)P(C|A)P(D|B)P(e|B)P(f|C,D)P(g|f)P(h|e,f)$

$= P(A)P(B)P(C|A)P(D|B) \sum_e \sum_f \sum_g \sum_h P(e|B)P(f|C,D)P(g|f)P(h|e,f)$

$= P(A)P(B)P(C|A)P(D|B) \sum_f P(f|C,D) \sum_g P(g|f) \sum_e P(e|B) \sum_h P(h|e,f)$

$= P(A)P(B)P(C|A)P(D|B) \underbrace{\sum_f P(f|C,D)}_{1} \underbrace{\sum_g P(g|f)}_{1} \underbrace{\sum_e P(e|B)}_{1} \underbrace{\sum_h P(h|e,f)}_{1}$

$= P(A)P(B)P(C|A)P(D|B)$

(b) Given what we learned about inference, simplify the probability P(A,D,E,H) using the CPTs by pushing in the summations and summing to 1 when possible.

P(A,D,E,H) =

$\sum_b \sum_c \sum_f \sum_g P(A)P(b)P(c|A)P(D|b)P(E|b)P(f|c,D)P(g|f)P(H|E,f) =$

$P(A)\sum_c P(c|A)\sum_f P(f|c,D)P(H|E,f)\sum_b P(b)P(D|b)P(E|b)\sum_g P(g|f) =$

$P(A)\sum_c P(c|A)\sum_f P(f|c,D)P(H|E,f)\sum_b P(b)P(D|b)P(E|b)\cancel{\sum_g P(g|f)}^{1} =$

$P(A)\sum_c P(c|A)\sum_f P(f|c,D)P(H|E,f)\sum_b P(b)P(D|b)P(E|b)$

(c) Starting from the summation in part (b), write the equation for each **factor** that is created and its size as you perform would variable elimination. Assume each variable is a binary variable.

Many possible answers depending on the order of your variable summations in part b. Below are the answers that correspond to my solution above.

Putting B as the innermost sum means that the first factor has 4 entries (for each combination of D and E). The second factor is of size 2^4 (D, E, and adding H and C). Leaving C as the outermost sum means that the variable A isn't in a factor table until the last factor (size 2^4). Switching the sum over C and the sum over F would result in the same size factors.

Note that if we had been given values of the variables and asked for exact probabilities, then we would have to normalize the last factor table.

Factor f_1 $f_1(D,E) = \sum_b P(b)P(D|b)P(E|b)$	**Size of** f_1 $2 * 2 = 4$
Factor f_2 $f_2(D,E,H,c) = \sum_f P(f|c,D)P(H|E,f)f_1(D,E)$	**Size of** f_2 $2 * 2 * 2 * 2 = 16$ (for each c, must have D,H,E)
Factor f_3 $f_3(A,D,E,H) = \sum_c P(c|A)f_2(D,E,H,c)$	**Size of** f_3 $2 * 2 * 2 * 2 = 16$
Factor f_4 None. We summed out g to 1	**Size of** f_4 None
Any final products $P(A,D,E,H) = P(A)f_3(A,D,E,H)$	**Final table size** 2^4