

# CSC412 Notes Week 4

Jerry Zheng

April 2021

## 1 Exact Inference

Lets say we have a distribution.  $P(x, y)$  If we want to perform inference on it we would use  $P(y|x) = \frac{P(x,y)}{\sum_y p(x,y)}$  However, there may be a set of variables in our model  $P$  that isn't apart of  $x$  or  $y$ .

$x$  = The observed evidence

$y$  = The unobserved variable we want to infer

$r = X - x, y$

Where  $r$  is a set of random variables neither apart of the query nor the evidence.

$$p(y|x) = \frac{p(y, x)}{p(x)}$$

each of the distributions we need to compute can be computed by marginalizing over the other variables.

$$p(y, x) = \sum_r p(y, x, r)$$

However, naively marginalizing over all unobserved variables requires a number of computations exponential in the number of random variables,  $(N)$ , in our model.

## 2 Variable Elimination

To compute this efficiently we will use the Variable Elimination Algorithm.

It's an exact inference algorithm, meaning it will calculate exactly  $p(y|x)$ .

It's also general, meaning it can be used on many different kinds of graphical models.

It's complexity depends of the conditional independence of our model.

It's intuitively done with dynamic programming.

## 2.1 Chain Example

$$A \rightarrow B \rightarrow C \rightarrow D$$

To find  $P(D)$ , we have the variables

$$\begin{aligned} y &= \{D\} \\ x &= \{\} \\ r &= \{A, B, C\} \end{aligned}$$

$$\begin{aligned} P(y) &= \sum_r p(y, r) \\ \Rightarrow P(D) &= \sum_{A, B, C} p(A, B, C, D) \\ &= \sum_A \sum_B \sum_C p(A)p(B|A)p(C|B)p(D|C) \end{aligned}$$

This is exponential in the number of variables  $\mathbf{O}(k^n)$  (k is the number of states per variable). But, reordering the joint distribution

$$P(D) = \sum_C p(D|C) \sum_B p(C|B) \sum_A p(A)p(B|A)$$

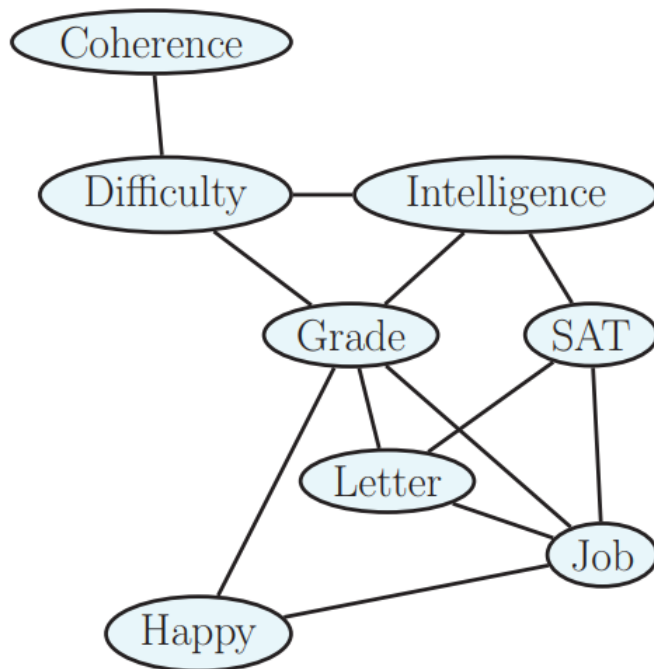
we can begin to simplify

$$\begin{aligned} P(D) &= \sum_C p(D|C) \sum_b p(C|B) \sum_A p(A)p(B|A) \\ &= \sum_C p(D|C) \sum_B p(C|B)\tau(B) \\ &= \sum_C p(D|C)\tau(C) \end{aligned}$$

So, by using dynamic programming to do the computation in reverse, we do inference over the joint distribution represented by the chain without generating it explicitly!

We have reduced the running time to  $\mathbf{O(nk^2)}$ !

## 2.2 bigger example



The joint distribution of our CS student graph is given by

$$P(C, D, I, G, S, L, J, H) = P(C)P(D|C)P(I)P(G|I, D)P(S|I)P(L|G)P(J|L, S)P(H|G, J)$$

with factors

$$P(C, D, I, G, S, L, J, H) = \psi_C(C), \psi_D(C, D), \psi_I(I), \psi_G(G, I, D), \psi_S(S, I), \psi_L(L, G), \psi_J(J, L, S), \psi_H(H, G, J)$$

(the textbook uses an undirected graph, not too important as variable elimination can be done all the same.)

(an explanation of  $\psi$  will be given in week 5's notes)

To compute  $p(J = 1)$ , we could calculate all possible assignments

$$p(J) = \sum_L \sum_S \sum_G \sum_H \sum_I \sum_D \sum_C p(C, D, I, G, S, L, J, H)$$

But we can do better with variable elimination. Where we push sums inside products.

$$\begin{aligned}
p(J) &= \sum_{L,S,G,H,I,D,C} p(C,D,I,G,S,L,J,H) \\
&= \sum_{L,S,G,H,I,D,C} \psi_C(C)\psi_D(D,C)\psi_I(I)\psi_G(G,I,D)\psi_S(S,I)\psi_L(L,G) \times \psi_J(J,L,S)\psi_H(H,G,J) \\
&= \sum_{L,S} \psi_J(J,L,S) \sum_G \psi_L(L,G) \sum_H \psi_H(H,G,J) \sum_I \psi_S(S,I)\psi_I(I) \times \sum_D \psi_G(G,I,D) \sum_C \psi_C(C)\psi_D(D,C)
\end{aligned}$$

from here we will marginalize out each variable individually to get a new factor at each step.

We do it in the order C, D, I, H, G, S, L to get P(J)

---


$$\begin{aligned}
&\sum_L \sum_S \psi_J(J,L,S) \sum_G \psi_L(L,G) \sum_H \psi_H(H,G,J) \sum_I \psi_S(S,I)\psi_I(I) \underbrace{\sum_D \psi_G(G,I,D) \sum_C \psi_C(C)\psi_D(D,C)}_{\tau_1(D)} \\
&\quad \sum_L \sum_S \psi_J(J,L,S) \sum_G \psi_L(L,G) \sum_H \psi_H(H,G,J) \sum_I \psi_S(S,I)\psi_I(I) \underbrace{\sum_D \psi_G(G,I,D)\tau_1(D)}_{\tau_2(G,I)} \\
&\quad \quad \sum_L \sum_S \psi_J(J,L,S) \sum_G \psi_L(L,G) \sum_H \psi_H(H,G,J) \underbrace{\sum_I \psi_S(S,I)\psi_I(I)\tau_2(G,I)}_{\tau_3(G,S)} \\
&\quad \quad \quad \sum_L \sum_S \psi_J(J,L,S) \sum_G \psi_L(L,G) \underbrace{\sum_H \psi_H(H,G,J) \tau_3(G,S)}_{\tau_4(G,J)} \\
&\quad \quad \quad \quad \sum_L \sum_S \psi_J(J,L,S) \underbrace{\sum_G \psi_L(L,G)\tau_4(G,J)\tau_3(G,S)}_{\tau_5(J,L,S)} \\
&\quad \quad \quad \quad \quad \underbrace{\sum_L \sum_S \psi_J(J,L,S)\tau_5(J,L,S)}_{\tau_6(J,L)} \\
&\quad \quad \quad \quad \quad \quad \underbrace{\sum_L \tau_6(J,L)}_{\tau_7(J)}
\end{aligned}$$