Exact Inference: Introduction

1

Exact Inference: Introduction

- Using a Bayesian network to compute probabilities is called inference
- In general, inference involves queries of the form:

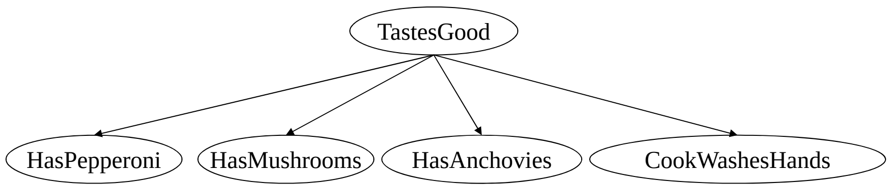
  P(X | **E**=**e**)

  E = The evidence variable(s)

  X = The query variable(s) (Assume a single variable for now)

2

Exact Inference: Introduction



- An example of a query would be:
  P(*TastesGood* = *true* | *HasPepperoni* = *true*,
  HasMushrooms = *true*, *HasAnchovies* = *false*)
- Note: Even though *CookWashesHands* is in the Bayesian network, it is not given values in the query (ie. they do not appear either as query variables or evidence variables)
- They are treated as unobserved variables

3

Exact Inference: Introduction

Recall that:

$$\mathbf{P}(X \mid \mathbf{E}=\mathbf{e}) = \alpha \mathbf{P}(X, \mathbf{E}=\mathbf{e})$$
$$= \alpha \sum_{\mathbf{y}} \mathbf{P}(X, \mathbf{E}=\mathbf{e}, \mathbf{Y}=\mathbf{y})$$

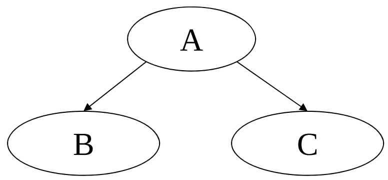
and $P(X_1, \ldots, X_n) = \prod_{i=1}^{n} P(X_i \mid parents(X_i))$

Enumeration-Ask algorithm:

Answer queries by computing sums of products of conditional probabilities from the network

4

Exact Inference: Introduction

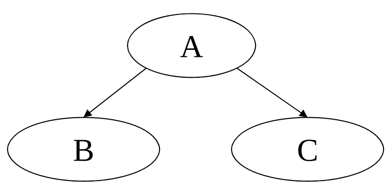


Query: P(A=true | B=true)

How do you solve this? 2 steps:

1. Express it in terms of the joint probability distribution P(A, B, C)
2. Express the joint probability distribution in terms of the entries in the CPTs of the Bayes net

Exact Inference: Introduction

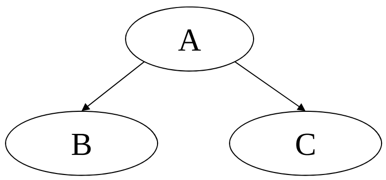


Whenever you see a conditional like P(A=true | B=true), use the Chain Rule:

P(A | B) = P(A, B) / P(B)

$$P(A = true \mid B = true) = \frac{P(A = true, B = true)}{P(B = true)}$$

Exact Inference: Introduction



Whenever you need to get a subset of the variables eg. P(B,A) from the full joint distribution P(A,B,C), use marginalization:

$$P(X) = \sum_{y} P(X, Y = y)$$

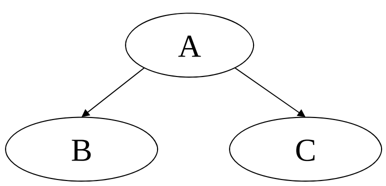
$$P(A = true \mid B = true)$$

$$= \frac{P(A = true, B = true)}{P(B = true)}$$

$$= \frac{\sum_{c} P(A = true, B = true, C = c)}{\sum_{a,c} P(A = a, B = true, C = c)}$$

7

Exact Inference: Introduction



To express the joint probability distribution as the entries in the CPTs, use:

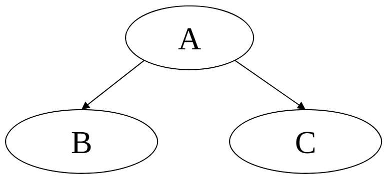
$$P(X_1, \ldots, X_N) = \prod_{i=1}^{N} P(X_i \mid Parents(X_i))$$

$$= \frac{\sum_{c} P(A = true, B = true, C = c)}{\sum_{a,c} P(A = a, B = true, C = c)}$$

$$= \frac{\sum_{c} P(C = c \mid A = true) P(B = true \mid A = true) P(A = true)}{\sum_{a,c} P(C = c \mid A = a) P(B = true \mid A = a) P(A = a)}$$

8

Exact Inference: Introduction



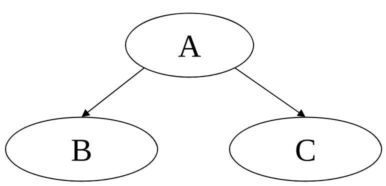
Take the probabilities that don't depend on the terms in the summation and move them outside the summation

$$= \frac{\sum_{c} P(C=c \mid A=true)P(B=true \mid A=true)P(A=true)}{\sum_{a,c} P(C=c \mid A=a)P(B=true \mid A=a)P(A=a)}$$

$$= \frac{P(B=true \mid A=true)P(A=true)\sum_{c} P(C=c \mid A=true)}{\sum_{a} P(B=true \mid A=a)P(A=a)\sum_{c} P(C=c \mid A=a)}$$

9

Exact Inference: Introduction



Simplify if possible.

$$= \frac{P(B=true \mid A=true)P(A=true)\sum_{c} P(C=c \mid A=true)}{\sum_{a} P(B=true \mid A=a)P(A=a)\sum_{c} P(C=c \mid A=a)}$$

$$= \frac{P(B=true \mid A=true)P(A=true)}{\sum_{a} P(B=true \mid A=a)P(A=a)}$$

10

Exact Inference: Introduction

Exact Inference in graphical models is NP-hard

– Exponential time in worst case

Approximate inference is also NP-hard

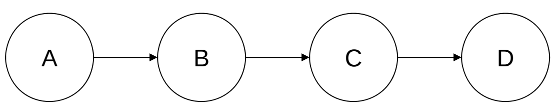
– But this is in the worst case. In practice, it is much more efficient

11

Exact Inference: Introduction

Example #2 (Variable Elimination):

$$P(B) = \sum_a P(A=a)P(B \mid A=a)$$



- Note: B is not instantiated with a value. We are computing the table *P(B)*.
- If *A* has *k* values and *B* has *k* values, the number of arithmetic operations required is $O(k^2)$
- If the chain has *n* nodes, computing the joint probability $P(X_1, \ldots, X_n)$ is $O(nk^2)$
- Naïve approach required $O(k^n)$ operations

12

Exact Inference: Introduction

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

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

Use dynamic programming to work from the innermost summation outward.

13

Exact Inference: Introduction

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

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

$$= \sum_C P(D \mid C) \sum_B P(C \mid B)\tau_1(B)$$

$$= \sum_C P(D \mid C)\tau_2(C)$$

$$\psi_1(A,B) = P(A)P(B \mid A)$$
$$\tau_1(B) = \sum_A \psi_1(A,B)$$

$$\psi_2(B,C) = \tau_1(B)P(C \mid B)$$
$$\tau_2(C) = \sum_B \psi_2(B,C)$$

14

Exact Inference: Introduction

Two key ideas to variable elimination:

1. Due to structure of BN, some subexpressions in the joint only depend on a small number of variables
2. Dynamic programming caches the intermediate results to avoid recomputing them exponentially many times

15