Reasoning Under Uncertainty: Bnet Inference

(Variable elimination)

CPSC 322 Lecture 28

Lecture Overview

- **Recap Learning Goals previous lecture**
- Bnets Inference
  - Intro
  - Factors
  - Variable elimination Intro

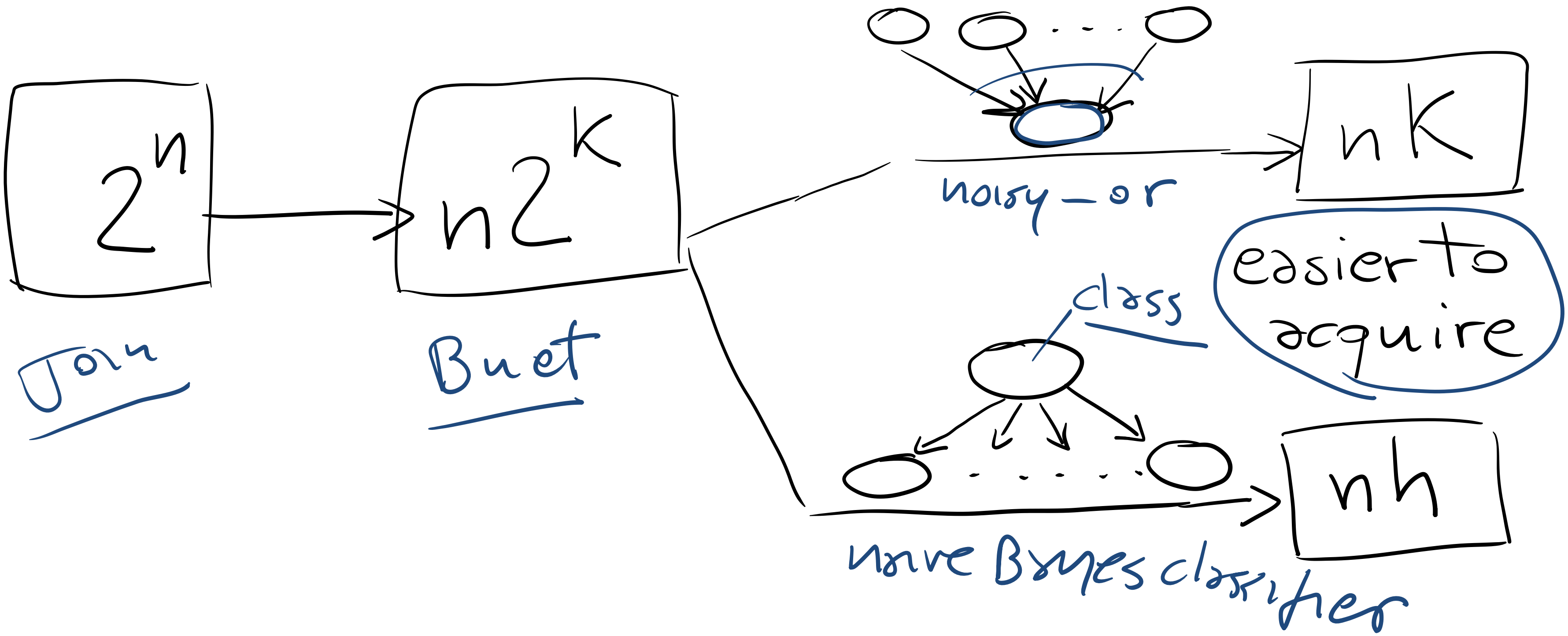
Learning Goals for previous class

You can:

- In a Belief Net, determine whether one variable is independent of another variable, given a set of observations.
- Define and use **Noisy-OR** distributions. Explain assumptions and benefit.
- Implement and use a **naïve Bayesian classifier**. Explain assumptions and benefit.

Bnets: Compact Representations

n Boolean variables, k max. number of parents



Only one parent with h possible values

Learning Goals for today's class

You can:

- Define **factors**. Derive new factors from existing factors. Apply operations to factors, including assigning, summing out and multiplying factors.

- Carry out **variable elimination** by using factor representation and using the factor operations. Use techniques to simplify variable elimination.

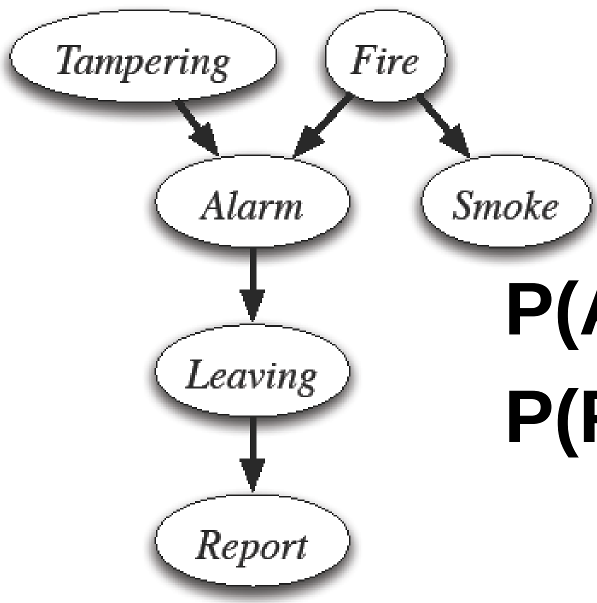
Lecture Overview

- Recap Learning Goals previous lecture
- **Bnets Inference**
  - Intro
  - Factors
  - Variable elimination Algo

Bnet Inference

- **Our goal:** compute probabilities of variables in a belief network

  What is the posterior distribution over **one** ~~or more variables~~, conditioned on one or more observed variables?



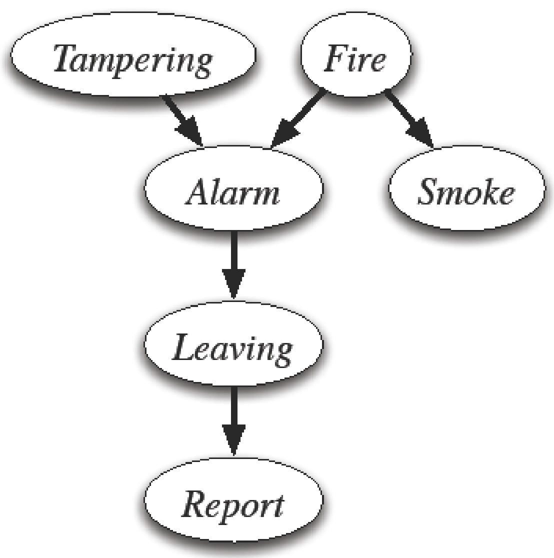
P(Alarm| Smoke = F)

P(Fire | Smoke = T, Leaving = F)

Bnet Inference: General

- Suppose the variables of the belief network are $X_1,\ldots,X_n$.
- Z is the **query variable**
- *$Y_1=v_1$, …, $Y_j=v_j$ are the **observed variables** (with their values)*
- $Z_1, \ldots, Z_k$ are the remaining variables

- What we want to compute: $P(Z \mid Y_1 = v_1 \ \ldots \ Y_j = v_j)$



Example:

P(L | S = t , R = f)

What do we need to compute?

Remember conditioning and marginalization…

$$P(L \mid S = t, R = f) = \frac{P(L, S = t, R = f)}{P(S = t, R = f)}$$

L	*S*	*R*	*P(L, S=t, R=f)*
t	t	f	.3
f	t	f	.2

.3 + .2 = .5

L	*S*	*R*	*P(L | S=t, R=f)*
t	t	f	.6
f	t	f	.4

In general…..

$$P(Z \mid Y_1 = v_1 \ \ldots \ Y_j = v_j) = \frac{P(Z, Y_1 = v_1 \ \ldots \ Y_j = v_j)}{P(Y_1 = v_1 \ \ldots \ Y_j = v_j)} = \frac{P(Z, Y_1 = v_1 \ \ldots \ Y_j = v_j)}{\sum_Z P(Z, Y_1 = v_1 \ \ldots \ Y_j = v_j)}$$

- We only need to **compute the numerator** and then **normalize**
- This can be framed in terms of **operations between factors** (that satisfy the semantics of probability)

Lecture Overview

- Recap Bnets

- Bnets Inference
  - Intro
  - **Factors**
  - Variable elimination Algo

Factors

A **factor** is a representation of a function from a tuple of random variables into a number.

- We will write factor *f* on variables $X_1, \ldots, X_j$ as $f(X_1, \ldots, X_j)$

A **factor** can denote:

- One distribution
- One *partial* distribution
- Several distributions
- Several *partial* distributions

over the given tuple of variables

Factor: Examples

$P(X_1, X_2)$ is a factor $f(X_1, X_2)$

Distribution

X_1	X_2	$f(X_1 , X_2)$
T	T	.12
T	F	.08
F	T	.08
F	F	.72

$P(X_1, X_2 = F)$ is a factor $f(X_1)_{X2=F}$

Partial distribution

X_1	X_2	$f(X_1)_{X2=F}$
T	F	.08
F	F	.72

Factors: More Examples

- A factor denotes one or more (possibly partial) distributions over the given tuple of variables
  - e.g., $P(X_1, X_2)$ is a factor $f(X_1, X_2)$ *Distribution*
  - e.g., $P(X_1, X_2, X_3 = v_3)$ is a factor $f(X_1, X_2)_{X3 = v3}$ *Partial distribution*
  - e.g., $P(X \mid Z,Y)$ is a factor $f(X,Z,Y)$ *Set of Distributions*
  - e.g., $P(X_1, X_3 = v_3 \mid X_2)$ is a factor $f(X_1, X_2)_{X3 = v3}$ *Set of partial Distributions*

iclicker.

f(X,Y,Z) ??

X	Y	Z	val
t	t	t	0.1
t	t	f	0.9
t	f	t	0.2
t	f	f	0.8
f	t	t	0.4
f	t	f	0.6
f	f	t	0.3
f	f	f	0.7

A. P(X,Y,Z)

B. P(Y|Z,X)

C. P(Z|X,Y)

D. None of the above

E. I want ice cream

Operations on factors

- Assigning values to variables
- Summing out variables
- Multiplying factors

Manipulating Factors

We can make new factors out of an existing factor

- Our first operation: we can *assign* some or all of the variables of a factor.

f(X,Y,Z):

X	Y	Z	val
t	t	t	0.1
t	t	f	0.9
t	f	t	0.2
t	f	f	0.8
~~f~~	~~t~~	~~t~~	~~0.4~~
~~f~~	~~t~~	~~f~~	~~0.6~~
~~f~~	~~f~~	~~t~~	~~0.3~~
~~f~~	~~f~~	~~f~~	~~0.7~~

What is the result of assigning X= t ?

f(X=t,Y,Z)

Or $f(X, Y, Z)_{X = t}$

Or f(Y,Z)

- Assignment **reduces** the factor **dimension**

(i.e. # of variables in the factor)

More examples of assignment

f(X,Y,Z):

X	Y	Z	val
t	t	t	0.1
t	t	f	0.9
t	f	t	0.2
t	f	f	0.8
f	t	t	0.4
f	t	f	0.6
f	f	t	0.3
f	f	f	0.7

f(X=t,Y,Z):

Y	Z	val
t	t	0.1
t	f	0.9
f	t	0.2
f	f	0.8

f(X=t,Y,Z=f):

Y	val
t	
f	

f(X=t,Y=f,Z=f):

val

Summing out a variable example

Our second operation: we can ***sum out*** a variable, say X_1 with domain $\{v_1, \ldots, v_k\}$, from factor $f(X_1, \ldots, X_j)$, resulting in a factor on $X_2, \ldots, X_j$ defined by:

$f_3(A,B,C)$:

B	A	C	val
t	t	t	0.03
t	t	f	0.07
f	t	t	0.54
f	t	f	0.36
t	f	t	0.06
t	f	f	0.14
f	f	t	0.48
f	f	f	0.32

$\sum_B f_3(A,B,C)$:

A	C	val
t	t	0.57
t	f	0.43
f	t	
f	f	

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

Multiplying factors

Our third operation: factors can be ***multiplied*** together.

$f_1(A,B)$:

A	B	Val
t	t	0.1
t	f	0.9
f	t	0.2
f	f	0.8

$f_2(B,C)$:

B	C	Val
t	t	0.3
t	f	0.7
f	t	0.6
f	f	0.4

$f_1(A,B) \times f_2(B,C)$:

A	B	C	val
t	t	t	
t	t	f	
t	f	t	
t	f	f	
f	t	t	
f	t	f	
f	f	t	
f	f	f	

Multiplying factors

Our third operation: factors can be ***multiplied*** together.

$f_1(A,B)$:

A	B	Val
t	t	0.1
t	f	0.9
f	t	0.2
f	f	0.8

$f_2(B,C)$:

B	C	Val
t	t	0.3
t	f	0.7
f	t	0.6
f	f	0.4

$f_1(A,B) \times f_2(B,C)$:

A	B	C	val
t	t	t	
t	t	f	
t	**f**	**t**	**??**
t	f	f	
f	t	t	
f	t	f	
f	f	t	
f	f	f	

A. 0.32

B. 0.54

C. 0.24

D. 0.06

E. I just feel like clicking E today…

Multiplying factors: Formal

- The **product** of factor $f_1(A, B)$ and $f_2(B, C)$, where B is the variable in common, is the factor $(f_1 \times f_2)(A, B, C)$ defined by:

$$f_1(A,B)\, f_2(B,C) = (f_1 \times f_2)(A,B,C)$$

Note1: it's defined on all A, B, C triples, obtained by multiplying together the appropriate pair of entries from f_1 and f_2.

Note2: A, B, C can be sets of variables

Factors Summary

- A factor is a representation of a function from a tuple of random variables into a number.
  - $f(X_1, \dots, X_j)$.
- We have defined three operations on factors:

1. Assigning one or more variables
   - $f(X_1=v_1, X_2, \dots, X_j)$ is a factor on $X_2, \dots, X_j$, also written as $f(X_1, \dots, X_j)_{X_1=v_1}$

2. Summing out variables is a factor on $X_2, \dots, X_j$
   - $\sum_{X_1} f(X_1, X_2, \dots, X_j) = f(X_1=v_1, X_2, \dots, X_j) + \dots + f(X_1=v_k, X_2, \dots, X_j)$

3. Multiplying factors
   - $f_1(A, B)\, f_2(B, C) = (f_1 \times f_2)(A, B, C)$

Lecture Overview

- **Recap Bnets**

- Bnets Inference
  - Intro
  - Factors
  - Intro to Variable Elimination

Variable Elimination Intro

- Suppose the variables of the belief network are $X_1,\ldots,X_n$.
- Z is the query variable
- *$Y_1=v_1$, …, $Y_j=v_j$ are the observed variables (with their values)*
- $Z_1, \ldots, Z_k$ are the remaining variables
- What we want to compute: $P(Z \mid Y_1 = v_1 \ \ldots \ Y_j = v_j)$
- We showed before that what we actually need to compute is

$$P(Z, Y_1 = v_1 \ \ldots \ Y_j = v_j)$$

This can be computed in terms of **operations between factors** (that satisfy the semantics of probability)

Variable Elimination Intro

- If we express the joint distribution as a single factor,

$$f\,(Z,\ \ Y_1\ldots,Y_j\,,\ \ Z_1\ldots,Z_j\ )$$

- We can compute $P(Z,Y_1{=}v_1,\ \ldots,Y_j{=}v_j)$ by
  - **assigning** $Y_1{=}v_1,\ \ldots,\ Y_j{=}v_j$
  - **summing out** the variables $Z_1, \ldots, Z_k$

$$P(Z,Y_1 = v_1, \ldots\ Y_j = v_j) = \sum_{Z_k} \ldots \sum_{Z_1} f\,(Z,Y_1,..,Y_j,Z_1,..,Z_k)_{Y_1=v_1,\ldots,Y_j=v_j}$$

Are we happy? NO, because the joint is too big to do this for large problems

Next Class

Variable Elimination

- The algorithm
- An example

Temporal models