Reasoning Under Uncertainty: Bnet Inference

(Variable elimination)

Computer Science cpsc322, Lecture 29

(Textbook Chpt 6.4)

Nov, 18, 2013

Lecture Overview

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

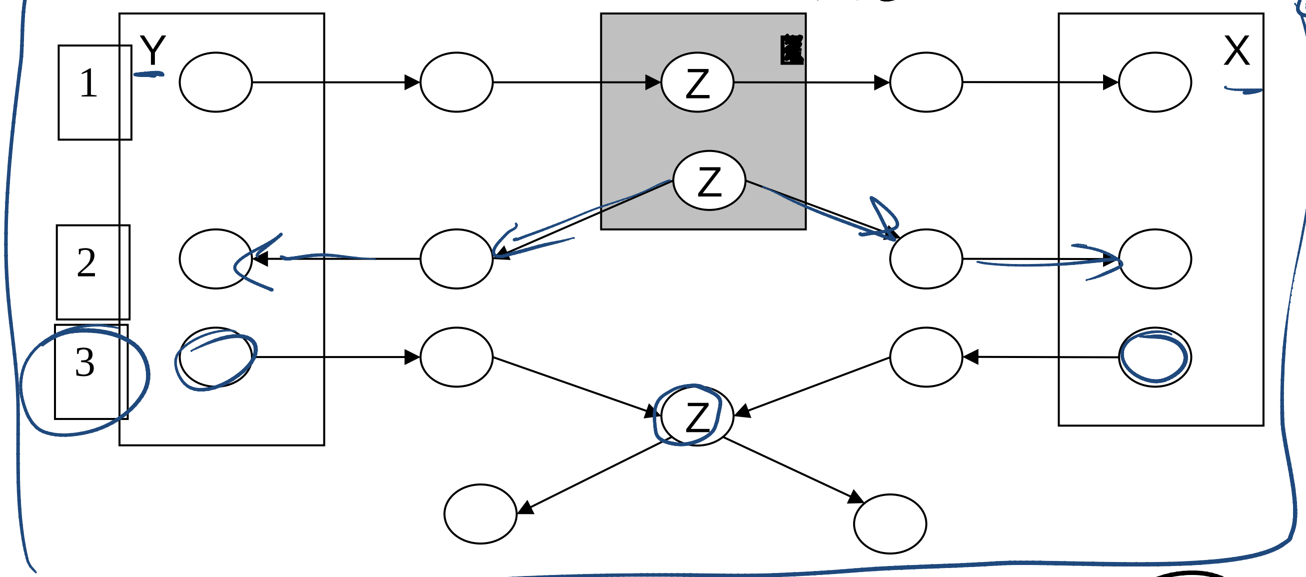
Learning Goals for Wed's class

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.

3 Configuration blocking dependency (belief propagation)

EVIDENCE/OBSERVED



true $\boxed{2}$

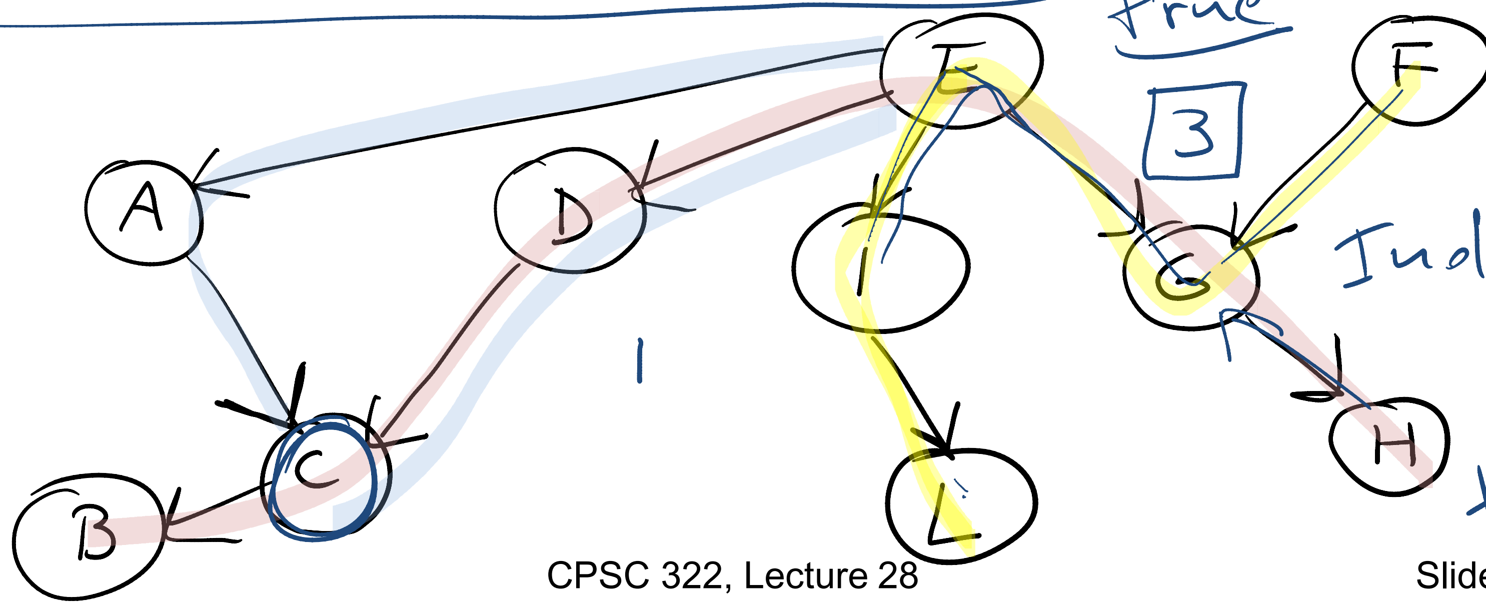
Indep(B,H {E})

false

Indep(E,C {D})

Indep(L,F {C})

true $\boxed{3}$

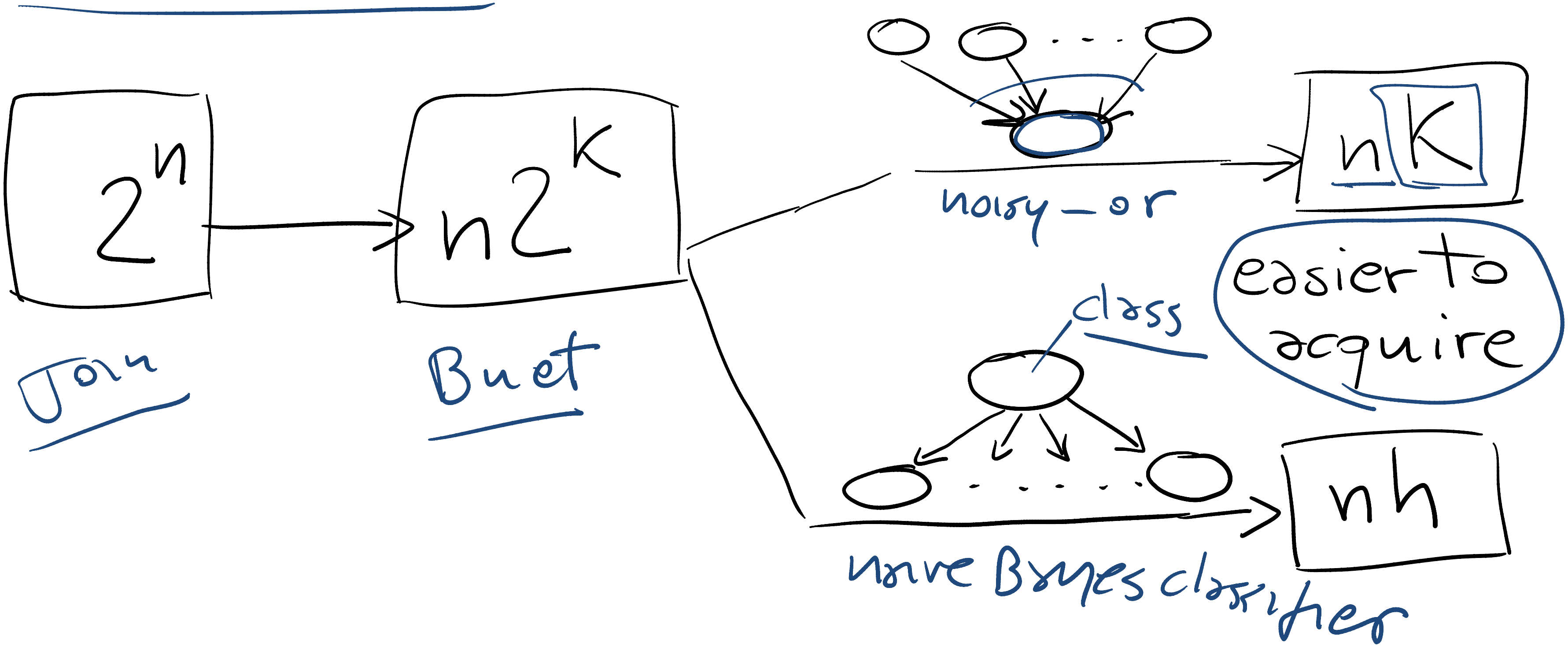


Ind(L,F {H})

false

Bnets: Compact Representations

n Boolean variables, k max. number of parents



Only one parent with h possible values

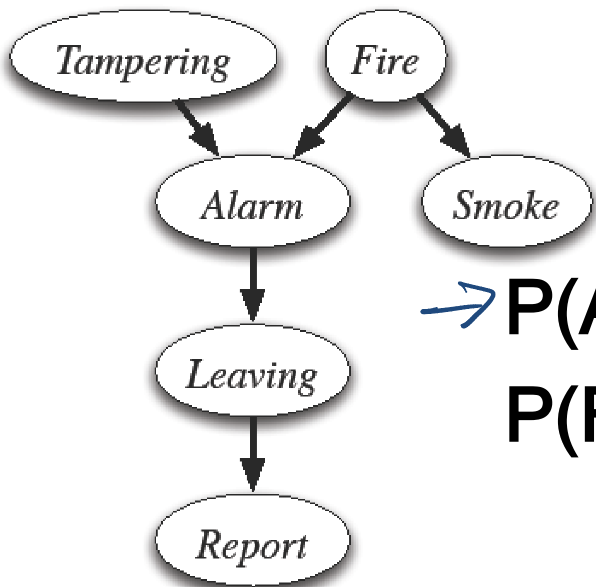
Lecture Overview

- 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?



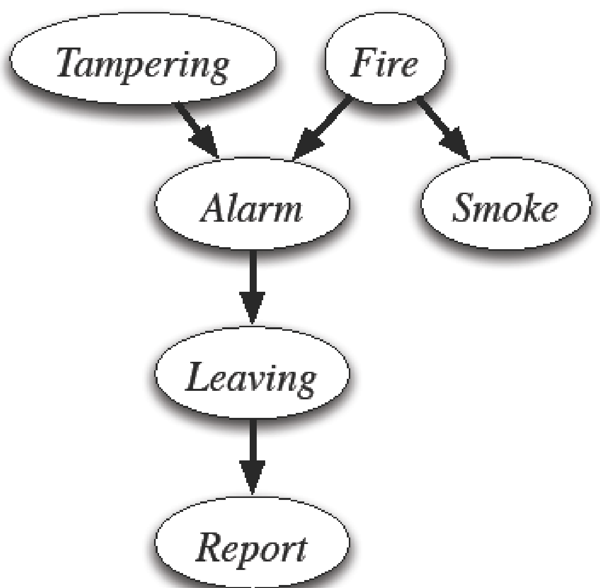
examples

→ **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
- *$Y_1=v_1, \ldots, Y_j=v_j$ are the observed variables (with their values)*
- $Z_1, \ldots, Z_k$ are the remaining variables
- What we want to compute: $\boxed{P(Z \mid Y_1 = v_1, \ldots, Y_j = v_j)}$



Example:

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

$Z \leftrightarrow L$

$Y_1\, Y_2 \leftrightarrow S, R$

$Z_1 Z_2 Z_3 \leftrightarrow T, F, A$

What do we need to compute?

Remember conditioning and marginalization…

3

$P(L \mid S = t, R = f) = \dfrac{P(L, S=t, R=f)}{P(S=t, R=f)}$ ← ① ②

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

iclicker

Do they have to sum up to one?

A. yes

B. no

② = .5

③

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

In general.....

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

1

2

- 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. $[0, 1]$
- 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

Partial distribution

$P(X_1, X_2 = v_2)$ is a factor $f(X_1)_{X2=v2}$

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.

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,Y,Z) ??

A. P(X,Y,Z)

B. P(Y|Z,X)

C. P(Z|X,Y)

D. None of the above

Factors

- A **factor** is a representation of a function from a tuple of random variables into a number. [0, 1]
- We will write factor f on variables $X_1,\dots,X_j$ as $f(X_1 \dots X_j)$
- A factor denotes one or more (possibly partial) distributions over the given tuple of variables

P(Z|X,Y)

- 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*

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

- 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*

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 $X=t$?

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

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

More examples of assignment

r(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

r(X=t,Y,Z):

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

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

Y	val
~~t~~	~~.9~~
f	.8

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

val
.8

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

$\Sigma_B f_3(A,C)$:

A	C	val
t	t	.57
t	f	.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	.03
t	t	f	.07
t	f	t	.054
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

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:

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

$$\begin{array}{ccc} & \begin{array}{cc} t & f \\ A & B \end{array} & \begin{array}{cc} f & f \\ B & C \end{array} \end{array}$$

t f f

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.
  - $f(X_1, \ldots, X_j)$.
- We have defined three operations on factors:

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

2. Summing out variables
   - $(\sum_{X_1} f)(X_2, \ldots, X_j) = f(X_1 = v_1, X_2, ,X_j) + \ldots + f(X_1 = v_k, X_2, ,X_j)$

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 Variable elimination Algo

Variable Elimination Intro

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

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

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

Variable Elimination Intro

- If we express the joint as a factor,

observed   sum out

$$f(Z, \underbrace{Y_1,\ldots,Y_j}_{\text{observed}}, Z_1,\ldots,Z_j)$$

assign

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

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

this is the Joint TOO BIG!

Are we done? no

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

Next Class

Variable Elimination

- The algorithm
- An example

Course Elements

- Work on Practice Exercises 6A and 6B
- Assignment 3 is due on Wed the 20^{th} !
- Assignment 4 will be available on Thur and due on Nov the 29^{th} (last class).