DS 2500 Mar 24

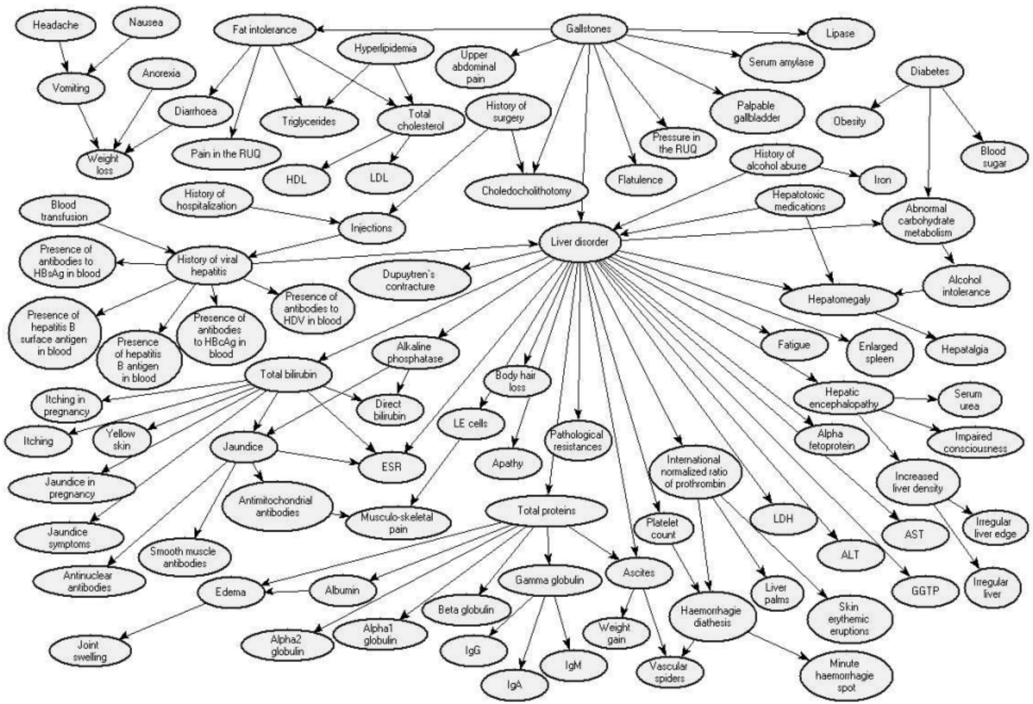
From last time (included for completeness today):

- bayes net motivation, definition

New:

- compute conditional probabilities with multiple random variables:
    - $P(ABC|XYZ) = P(ABCXYZ) / P(XYZ)$
- computing conditional probabilities
    - via spreadsheet ("computer" method)

WHAT ARE BAYES NETS GOOD FOR?



source: https://sites.pitt.edu/~druzdzel/psfiles/cbmi99a.pdf

Bayes nets allow us to incorporate multiple pieces of evidence into some conditional prob of interest:

given a person has:
- symptom 4
- symptom 11
- risk factor 7

whats the prob of liver disorder?

Bayesian Network (Bayes Net)

$$A \rightarrow B \rightarrow C$$

(formally):
A directed, acyclic graph which represents conditional distributions / independences between a set of random variables.

$$\begin{tikzcd} & B \arrow[d] \\ A \arrow[ur] & C \arrow[l] \end{tikzcd}$$

HAS CYCLE, NOT A BAYES NET

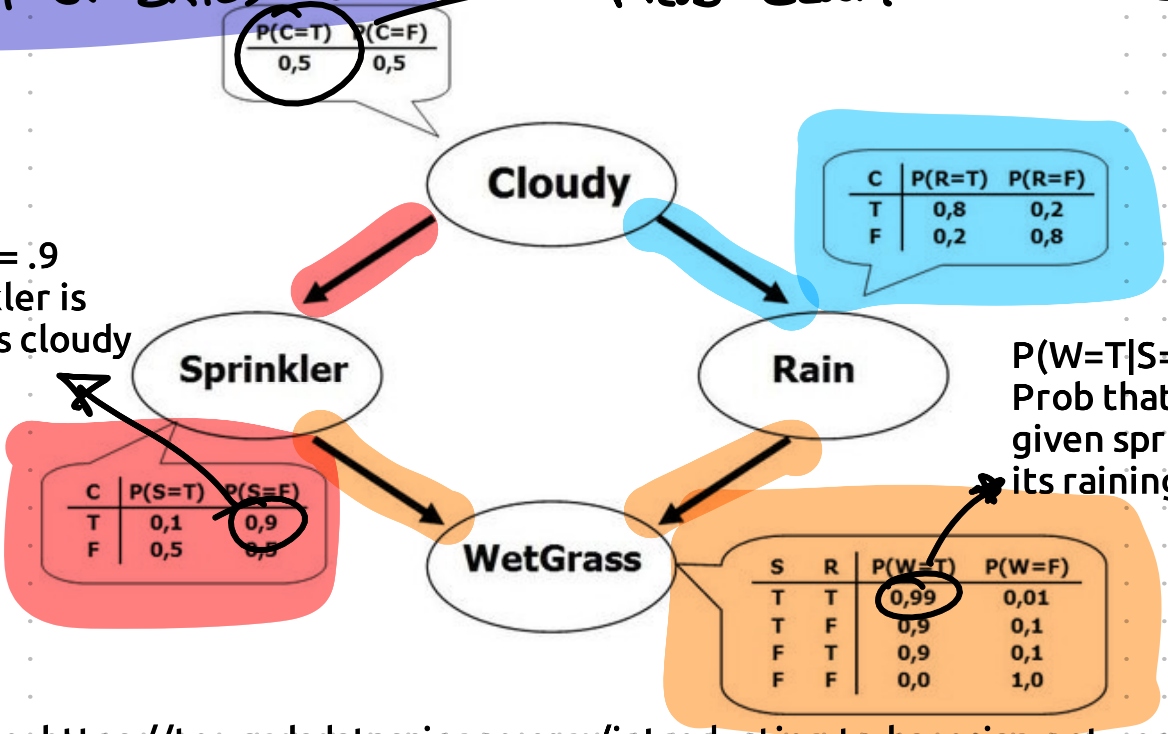
each node represents a random variable
directed edges represent conditional distributions
any node without inward edges has prob specified (its part of "bayes net" too!)

$$P(A) \quad A \xrightarrow[P(B|A)]{} B \xrightarrow[P(C|B)]{} C$$

(informally):
a network which describes how random variables influence each other. can be used to compute conditional probabilities of interest

Anatomy of Bayes Net

Prob Cloudy = True is 50%



P(S=F|C=T) = .9
Prob sprinkler is off given it's cloudy out is 90%

P(W=T|S=T R=T) = .99
Prob that grass is wet given sprinkler is on and its raining is 99%

source: https://towardsdatascience.com/introduction-to-bayesian-networks-81031eeed94e

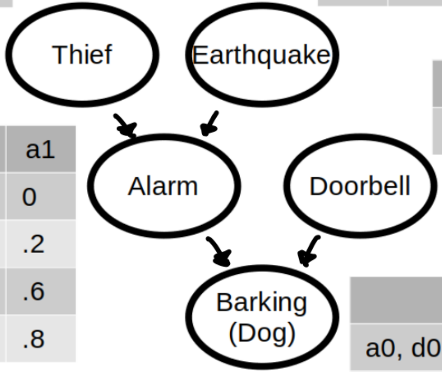
Bayes Net Notation (our convention)

Each random variable is denoted with a capital letter (T for Thief). Each outcome in sample space has its own lowercase letter :

t0 = no thief
t1 = thief

t0	t1
.99	.01

e0	e1
.95	.05



d0	d1
.8	.2

	a0	a1
t0, e0	1	0
t0, e1	.8	.2
t1, e0	.4	.6
t1, e1	.2	.8

	b0	b1
a0, d0	1	0
a0, d1	.2	.8
a1, d0	.5	.5
a1, d1	.01	.99

(quick) ICA X:

what's prob of earthquake?

given a thief in house, but no earthquake, what's prob alarm goes off?

interpretation question:
- is alarm better at detecting thieves or earthquakes?

- which sound bothers the dog more, the alarm or doorbell?

In Class Assignment (last time):
Estimate / intuite the four probabilities below. Is it greater / lesser / equal to other prob immediately above?

What is the prob of thief? $P(T=1) = 1\%$

Given that alarm is going off, what is prob of thief?

$$P(T=1 \mid A=1) > P(T=1)$$

THIEF IS MORE COMMON WHEN ALARM GOES OFF THAN NOT

Given that alarm is going off & dog is barking, what is prob of thief?

$$P(T=1 \mid A=1\ B=1) = P(T=1 \mid A=1)$$

GIVEN ALARM THIEF AND BARK ARE INDEPENDENT

Given that alarm is going off, dog is barking & earthquake, what is prob of thief?

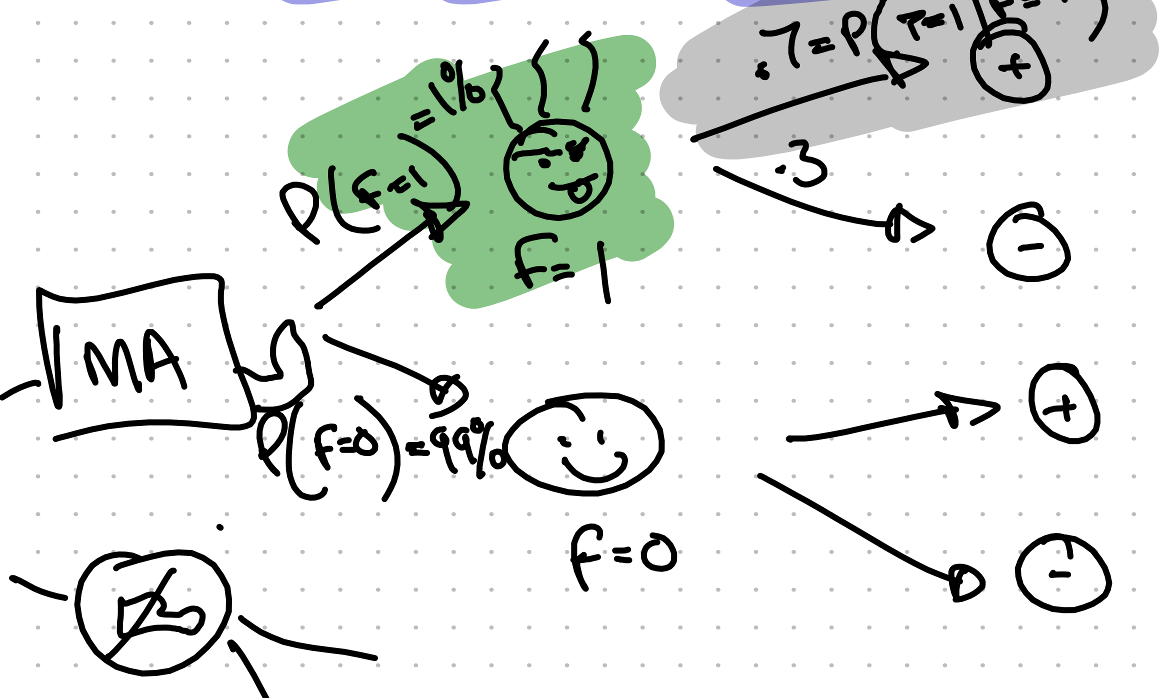
$$P(T=1 \mid A=1\ B=1\ E=1) < P(T=1 \mid A=1\ B=1)$$

EARTHQUAKE EXPLAINS AWAY ALARM

How do we compute conditional probabilities from a Bayes Net?

With a computer:
Step 1: Rewrite conditional probability without conditional
Step 2: In a spreadsheet, compute prob of every possible combination of outputs for all vars
Step 3: Compute the needed probabilities from step 1 via marginalization

$$P(a|b)P(b) = P(ab)$$



$$P(T=1\ F=1) = P(F=1)P(T=1|F=1)$$

$$P(T=1\ F=1|M=1) = P(F=1|M=1)\,P(T=1|F=1\ M=1)$$

$$P(ab) = P(a|b)P(b)$$

$$P(a|b) = \frac{P(ab)}{P(b)}$$

Step 1: write conditional probabilities as ratio of (not conditional) probabilities

$$P(ABC \mid XYZ) = \frac{P(ABC\,XYZ)}{P(XYZ)}$$

PROB OF TARGET AND EVIDENCE TOGETHER

PROB OF EVIDENCE

TARGET VARS

EVIDENCE VARS

(ex: Given that alarm is going off & dog is barking, what is prob of thief?)

$$P(t_1 \mid a_1 b_1) = \frac{P(t_1 a_1 b_1)}{P(a_1 b_1)}$$

Step 2: In a spreadsheet, compute prob of every possible combination of outputs for all vars

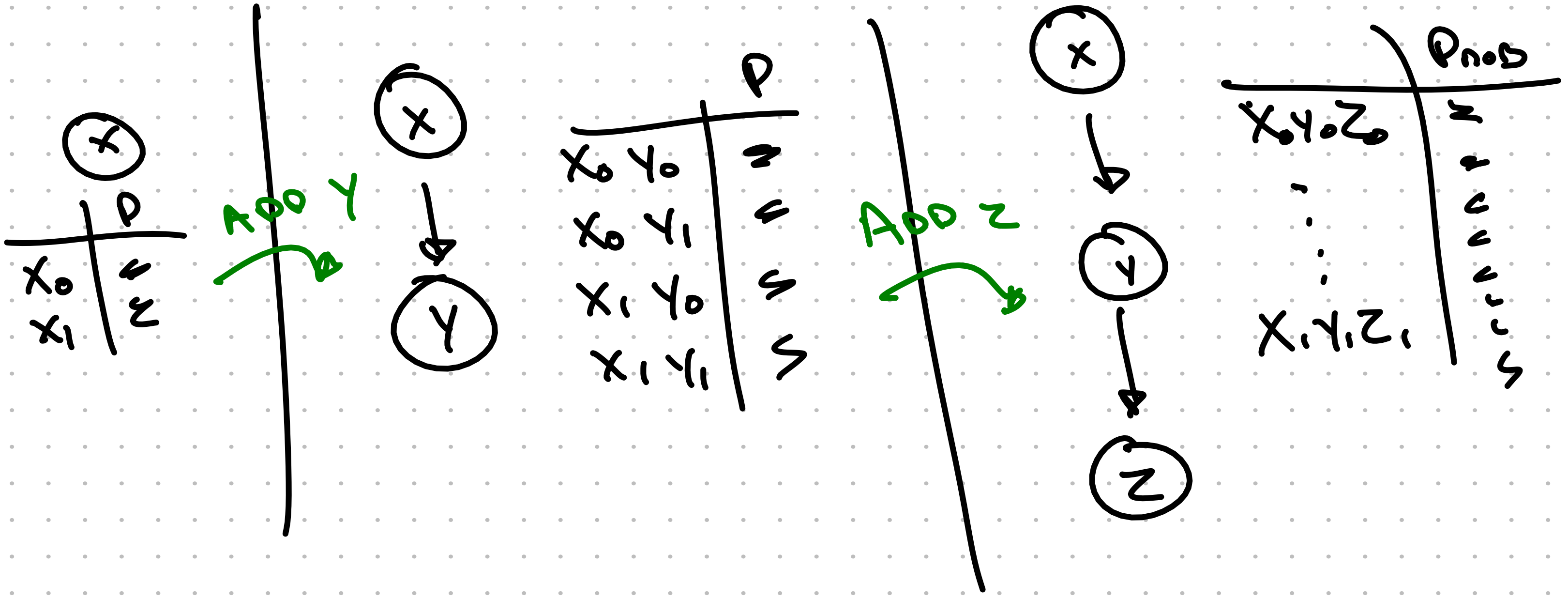
("JOINT DISTRIBUTION TABLE")

EVERY COMBINATION VARS

PROB

B: Barking	D: Doorbell	A: Alarm	T: Thief	E: Earthqua	P(BDATE)
b0	d0	a0	t0	e0	0.7524
b0	d0	a0	t0	e1	0.03168
b0	d0	a0	t1	e0	0.00304
b0	d0	a0	t1	e1	8E-05
b0	d0	a1	t0	e0	0
b0	d0	a1	t0	e1	0.00396
b0	d0	a1	t1	e0	0.00228
b0	d0	a1	t1	e1	0.00016
b0	d1	a0	t0	e0	0.03762
b0	d1	a0	t0	e1	0.001584
b0	d1	a0	t1	e0	0.000152
b0	d1	a0	t1	e1	4E-06
b0	d1	a1	t0	e0	0
b0	d1	a1	t0	e1	1.98E-05
b0	d1	a1	t1	e0	1.14E-05
b0	d1	a1	t1	e1	8E-07
b1	d0	a0	t0	e0	0
b1	d0	a0	t0	e1	0

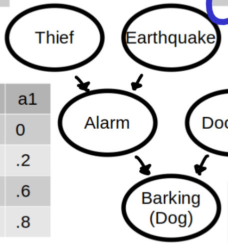
Producing a Joint Table Iteratively



Producing Joint Table (Adding independent Nodes)

t0	t1
.99	.01

e0	e1
.95	.05



d0	d1
.8	.2

	a0	a1
t0, e0	1	0
t0, e1	.8	.2
t1, e0	.4	.6
t1, e1	.2	.8

	b0	b1
a0, d0	1	0
a0, d1	.2	.8
a1, d0	.5	.5
a1, d1	.01	.99

Since T, E are independent.

$$P(TE) = P(T)P(E)$$

T	P(T)
t_0	.99
t_1	.01

T	E	P(TE)
t_0	e_0	.99 · .95
t_0	e_1	.99 .05
t_1	e_0	.01 · .95
t_1	e_1	.01 · .05

Producing Joint Table (Dependent Nodes)

$X \rightarrow Y$

	Y_0	Y_1
X_0	$1/3$	$2/3$
X_1	$1/4$	$3/4$

These are $P(Y|X)$ values

Since Y depends on X

$$P(X\,Y) = P(Y|X)\,P(X)$$

X	Prob
X_0	$1/7$
X_1	$6/7$

X	Y	$P(X\,Y)$
X_0	Y_0	$1/7 \cdot 1/3$
X_0	Y_1	$1/7 \cdot 2/3$
X_1	Y_0	$6/7 \cdot 1/4$
X_1	Y_1	$6/7 \cdot 3/4$

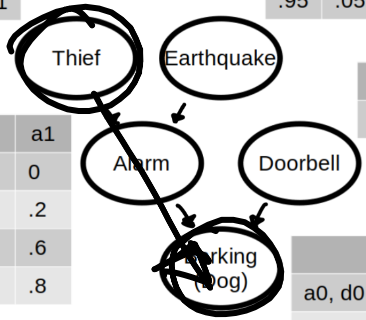
In Class Exercise (don't submit):

Build the joint distribution table for the bayes net on the left.

(You needn't submit for credit. You can check your work with the given final answer csv on website)

t0	t1
.99	.01

e0	e1
.95	.05



d0	d1
.8	.2

	a0	a1
t0, e0	1	0
t0, e1	.8	.2
t1, e0	.4	.6
t1, e1	.2	.8

	b0	b1
a0, d0	1	0
a0, d1	.2	.8
a1, d0	.5	.5
a1, d1	.01	.99

X	Y	Z	Prob
x_0	y_0	z_0	1/4
x_0	y_0	z_1	0
x_0	y_1	z_0	0
x_0	y_1	z_1	1/8
x_1	y_0	z_0	3/8
x_1	y_0	z_1	0
x_1	y_1	z_0	0
x_1	y_1	z_1	1/4

Marginalizing in Joint Table (step 3)

Compute $P(x_0 z_0)$

$$= P(x_0 z_0 y_0) + P(x_0 z_0 y_1)$$

$$= 1/4 + 0 = 1/4$$

Quick Practice

Compute $P(y_1 x_1)$

Compute $P(x_0)$

Marginalizing in Joint Table
(step 3)

X	Y	Z	Prob
x_0	y_0	z_0	1/4
x_0	y_0	z_1	0
x_0	y_1	z_0	0
x_0	y_1	z_1	1/8
x_1	y_0	z_0	3/8
x_1	y_0	z_1	0
x_1	y_1	z_0	0
x_1	y_1	z_1	1/4

Quick Practice

Compute $P(y_1, x_1) = 0 + 1/4 = 1/4$

Compute $P(x_0) = 1/4 + 1/8 = 3/8$

Putting it all together:

Step 1: Rewrite conditional probability without conditional
Step 2: In a spreadsheet, compute prob of every possible combination of outputs for all vars
Step 3: Compute the needed probabilities from step 1 via marginalization

Example:

Given alarm is going off and dog is barking, what is the probability of a thief?

$$P(t_1 \mid a_1 b_1) = \frac{P(t_1 a_1 b_1)}{P(a_1 b_1)} \approx .381$$

Putting it all together:

Step 1: Rewrite conditional probability without conditional
Step 2: In a spreadsheet, compute prob of every possible combination of outputs for all vars
Step 3: Compute the needed probabilities from step 1 via marginalization

Example:

Given alarm is going off and dog is barking, what is the probability of a thief?

$$P(t_1 \mid a_1 b_1 e_1) = \frac{P(t_1\, a_1 b_1 e_1)}{P(a_1 b_1 e_1)} \approx .03883$$

In Class Exercise 1:

Explicitly compute each of the following

1. What is the prob of thief? $P(t_1) = .01$

2. Given that alarm is going off, what is prob of thief? $P(t_1 | a_1) = .38125$

3. Given that alarm is going off & dog is barking, what is prob of thief? $P(t_1 | a_1 b_1) = .381$

4. Given that alarm is going off, dog is barking & earthquake, what is prob of thief? $P(t_1 | a_1 b_1 e_1) = .03885$

Answer each question below with one sentence (please avoid algebraic motivations and appeal to our intuition):

- Why is the prob of 2 greater than the prob of 1?
- Why is the prob of 3 equal to the prob of 2?
- Why is the prob of 4 less than the prob of 2?

$$P(t_1 | a_1) = \frac{P(a_1, t_1)}{P(a_1)} = \frac{.0061}{.016}$$

$$= .38125$$

That spreadsheet work sure was cumbersome ... if only we could make the computer do all that busy work!

In Class Assignment 2:

Design the interface of a Bayes Net "library" which allows the user to

1. specify a bayes net on discrete random variables (as shown here)

as well as querying the bayes net for arbitrary:

2. conditional distributions
3. marginal distributions (i.e. not conditional distributions)

This design is intentionally open-ended. You're welcome to write out notes on paper, though you can write function (or method) docstrings as well if you're ready.