Announcements

Assignments:

- HW9 (written)
  - Due Tuesday 3/31, 10 pm
- P4
  - Due Thu 4/2, 10 pm
  - Check your slip days!

Midterm:

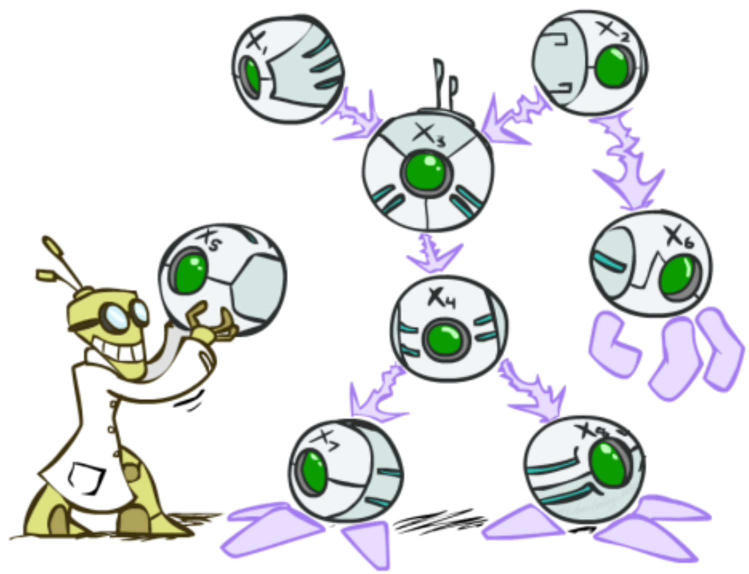
- Wed 4/8 on Gradescope (more details coming soon)
- Covers material from Integer Programming until Reinforcement Learning

Course Feedback:

- Don’t forget to fill out the 3 feedback forms on Piazza

AI: Representation and Problem Solving

Bayes Nets



Instructors: Pat Virtue & Stephanie Rosenthal

Slide credits: CMU AI and http://ai.berkeley.edu

Probability Notation

Notation and conventions in this course

$$P(+b, C) = \sum_{a \in \{a_1, a_2, a_3\}} P(a, +b, C)$$

Probability Notation

Notation and conventions in this course

$$P(+b, C) = \sum_{a \in \{a_1, a_2, a_3\}} P(a, +b, C)$$

- Random variables:
  - Capitalized
  - Represents all potential outcomes
  - e.g. C
- Outcomes (values):
  - lower case
  - e.g. $+b$, a_1, a_2, a_3
- Variables for values:
  - lower case
  - E.g. a

- For each random variable
  - Discrete outcomes
  - Disjoint outcomes
  - Not always binary

Probability Notation

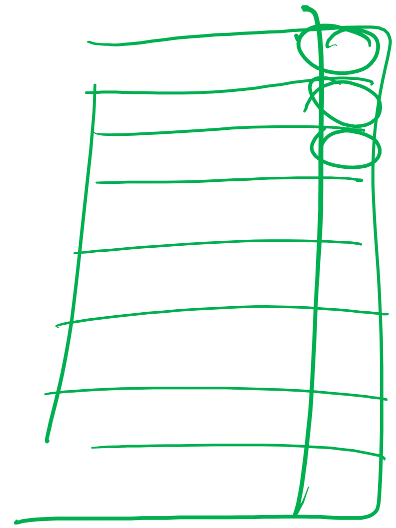
Notation and conventions in this course

$$P(+b, C) = \sum_{a \in \{a_1, a_2, a_3\}} P(a, +b, C)$$

$+c \quad -c$

$$= \sum_a P(a, \cancel{+}b, C)$$

$$= \sum_A P(A, +b, C)$$



$$P(A, B, C) = P(A)P(B|A)P(C|A,B)$$

$$P(+a, -b, +c)$$

$$P(+b, +c)$$

$$P(+b, -c)$$

Piazza Poll 1

PAUSE!

Which of the following probability tables sum to one?

Select all that apply.

i. $P(A \mid b)$

ii. $P(A, B, c)$

iii. $P(A, B \mid c)$

iv. $P(a, b \mid c)$

v. $P(a \mid B, C)$ ← Calamity

vi. $P(c \mid A)$

Piazza Poll 1

Which of the following probability tables sum to one?

Select all that apply.

i. $P(A \mid b)$

ii. $P(A, B, c)$

iii. $P(A, B \mid c)$

iv. $P(a, b \mid c)$

v. $P(a \mid B, C)$

vi. $P(c \mid A)$

Piazza Poll 1

Which of the following probability tables sum to one?

Select all that apply.

i. $P(A \mid b)$

ii. $P(A, B, c)$

iii. $P(A, B \mid c)$

iv. $P(a, b \mid c)$

v. $P(a \mid B, C)$

vi. $P(c \mid A)$

Piazza Poll 2

PAUSE!

How many valid equations can we compose using:

$$P(x),\ P(y),\ P(x,y),\ P(x|y),\ P(y|x) \text{ and } =,\ \times,\ \div$$

First one: $P(x|y) = P(x,y)/P(y)$

$P(y|x) = P(x,y)/P(x)$

Conditional probability

$P(x,y) = P(y|x)P(x)$

$P(x,y) = P(x|y)P(y)$

Product

$P(y|x)P(x) = P(x|y)P(y)$

$P(x|y) = \frac{P(y|x)P(x)}{P(y)}$

$P(y|x) = \frac{P(x|y)P(y)}{P(x)}$

Bayes

~~A) 2~~

B) 4

C) 7

D) Other

At most one use per probability term

e.g. Not $P(x) = P(x)$

Must be different

e.g. Cannot also use $P(x,y)/P(y) = P(x|y)$

Piazza Poll 2

How many valid equations can we compose using:

$P(X)$, $P(Y)$, $P(X,Y)$, $P(X|Y)$, $P(Y|X)$ and $=$, $\times$, $\div$

First one: $P(X|Y) = P(X,Y)/P(Y)$

At most one use per probability term

e.g. Not $P(X) = P(X)$

Must be different

e.g. Cannot also use $P(X,Y)/P(Y) = P(X|Y)$

A) 2

B) 4

C) 7

D) Other

Probability Tools Summary

Our toolbox

1. Definition of conditional probability $$P(A|B) = \frac{P(A,B)}{P(B)}$$

2. Product Rule $$P(A,B) = P(A|B)P(B)$$

3. Bayes' theorem $$P(B|A) = \frac{P(A|B)P(B)}{P(A)}$$

$P(A,B,C) = P(A)P(B|A)P(C|A,B)$

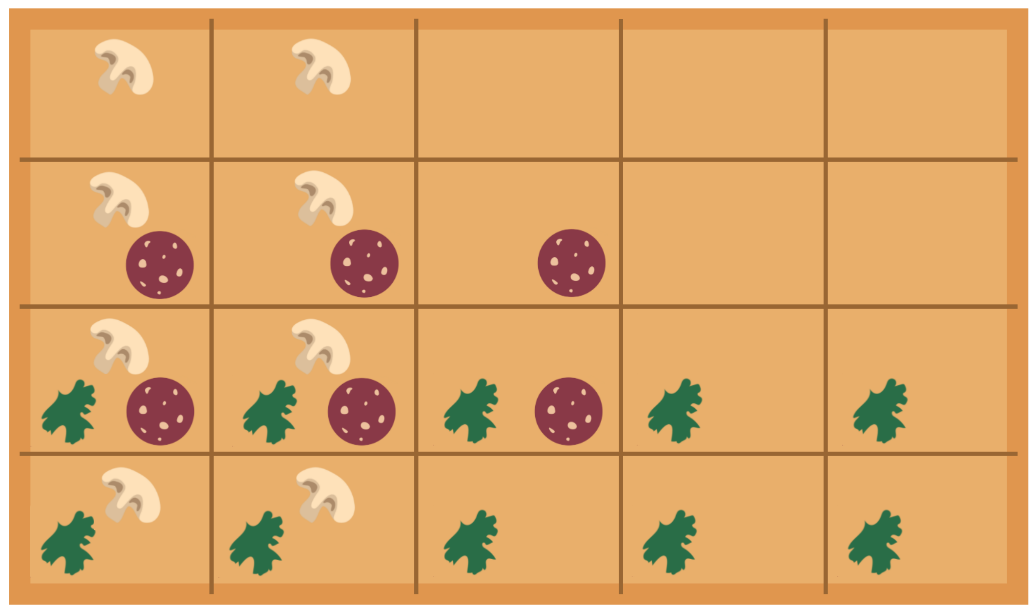
4. Chain Rule $$\underline{P(X_1, \dots, X_N)} = \prod_{n=1}^{N} P(X_n \mid X_1, \dots, X_{n-1})$$

Answer Any Query from Joint Distribution

What is the probability of getting a slice with:

1) No mushrooms
2) Spinach and no mushrooms
3) Spinach, when asking for slice with no mushrooms

- Mushrooms
- Spinach
- No spinach
- No spinach and mushrooms
- No spinach when asking for no mushrooms
- No spinach when asking for mushrooms
- Spinach when asking for mushrooms



Icons: CC, https://openclipart.org/detail/296791/pizza-slice

Answer Any Query from Joint Distribution

You can answer all of these questions:

$P(M)$

m_1	$12/20$
m_2	

$P(S)$

s_1	
s_2	

$P(M,S)$

m_1	s_1	
m_1	s_2	$6/20$
m_2	s_1	
m_2	s_2	

$P(M|\mathrm{s}_1)$

m_1	
m_2	

$P(M|\mathrm{s}_2)$

m_1	
m_2	

$P(S|m_1)$

s_1	
s_2	$6/12$

$P(S|\mathrm{m}_2)$

s_1	
s_2	

Answer Any Query from Joint Distribution

P(Weather)?

P(Weather | winter)?

P(Weather | winter, hot)?

Season	Temp	Weather	P(S, T, W)
summer	hot	sun	0.30
summer	hot	rain	0.05
summer	cold	sun	0.10
summer	cold	rain	0.05
winter	hot	sun	0.10
winter	hot	rain	0.05
winter	cold	sun	0.15
winter	cold	rain	0.20

Answer Any Query from Joint Distribution

P(Weather)?

$P(\text{Weather}=\text{sun}) =$

$.3 + .1 + .1 + .15 = .65$

$\sum_t \sum_s P(s, t, W=\text{sun})$

$P(\text{Weather}=\text{rain}) =$

Season	Temp	Weather	P(S, T, W)
summer	hot	sun	0.30
summer	hot	rain	0.05
summer	cold	sun	0.10
summer	cold	rain	0.05
winter	hot	sun	0.10
winter	hot	rain	0.05
winter	cold	sun	0.15
winter	cold	rain	0.20

Answer Any Query from Joint Distribution

P(Weather | winter)?

$$P(\text{Weather}=\text{sun} \mid \text{winter}) = \frac{.1 + .15}{.1 + .05 + .15 + .2} = .5$$

$$\frac{\sum_t P(S=\text{winter}, t, W=\text{sun})}{\sum_w \sum_t P(S=\text{winter}, t, w)} = \frac{P(\text{winter}, \text{sun})}{P(\text{winter})} = P(\text{sun} \mid \text{winter})$$

conditional probability

Season	Temp	Weather	P(S, T, W)
summer	hot	sun	0.30
summer	hot	rain	0.05
summer	cold	sun	0.10
summer	cold	rain	0.05
winter	hot	sun	0.10
winter	hot	rain	0.05
winter	cold	sun	0.15
winter	cold	rain	0.20

Answer Any Query from Joint Distribution

P(Weather | winter, hot)?

Season	Temp	Weather	P(S, T, W)
summer	hot	sun	0.30
summer	hot	rain	0.05
summer	cold	sun	0.10
summer	cold	rain	0.05
winter	hot	sun	0.10
winter	hot	rain	0.05
winter	cold	sun	0.15
winter	cold	rain	0.20

Answer Any Query from Joint Distribution

Two tools to go from joint to query

1. Definition of conditional probability

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

2. Law of total probability (marginalization, summing out)

$$P(A) = \sum_b P(A,b)$$

$$P(Y \mid U,V) = \sum_x \sum_z P(x,Y,z \mid U,V)$$

Answer Any Query from Joint Distribution

Two tools to go from joint to query

Joint: $P(H_1, H_2, Q, E)$

Query: $P(Q \mid e)$

1. Definition of conditional probability

$$P(Q|e) = \frac{P(Q, e)}{P(e)}$$

2. Law of total probability (marginalization, summing out)

$$P(Q, e) = \sum_{h_1} \sum_{h_2} P(h_1, h_2, Q, e)$$

$$P(e) = \sum_{q} \sum_{h_1} \sum_{h_2} P(h_1, h_2, q, e)$$

Answer Any Query from Joint Distribution

P(Weather)?

P(Weather | winter)?

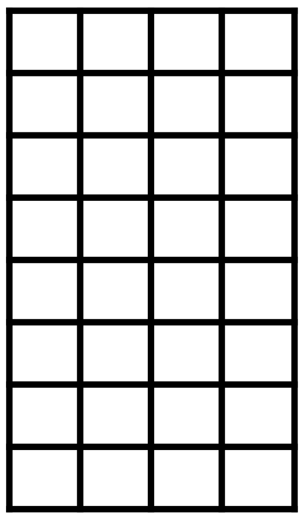
P(Weather | winter, hot)?

Season	Temp	Weather	P(S, T, W)
summer	hot	sun	0.30
summer	hot	rain	0.05
summer	cold	sun	0.10
summer	cold	rain	0.05
winter	hot	sun	0.10
winter	hot	rain	0.05
winter	cold	sun	0.15
winter	cold	rain	0.20

Answer Any Query from Joint Distribution

Joint distributions are the best!

Joint



Query

$$\Rightarrow \quad P(q_1, q_2 \mid e_1, e_2, e_3)$$

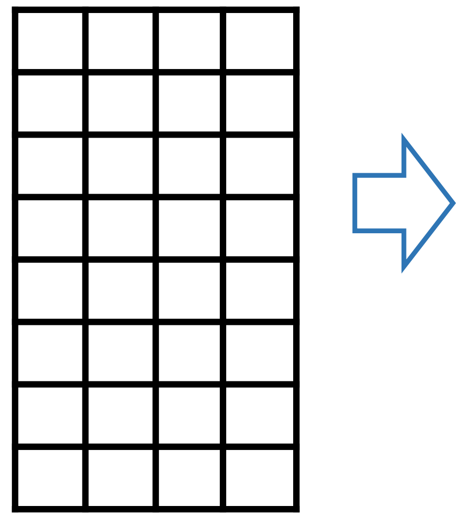
Answer Any Query from Joint Distribution

Joint distributions are the best!

Problems with joints

- We aren’t given the joint table
  - Usually some set of conditional probability tables

Joint



Query

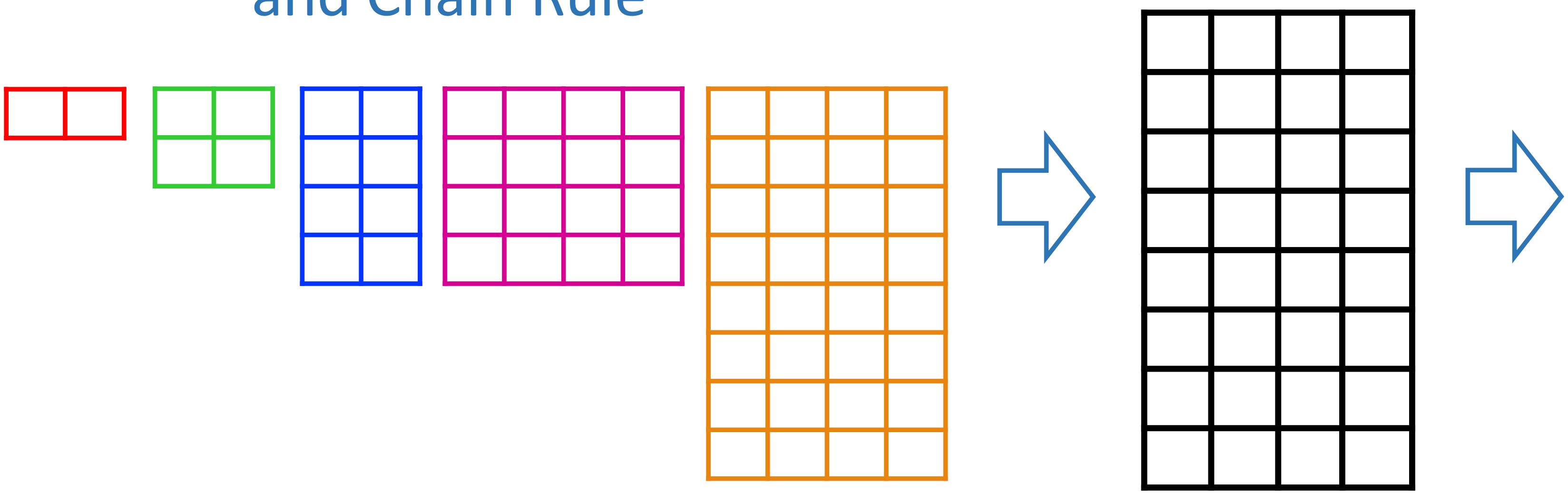
$P(a \mid e)$

Build Joint Distribution Using Chain Rule

Conditional Probability Tables and Chain Rule

Joint

Query



$P(a \mid e)$

$P(A)\, P(B \mid A)\, P(C \mid A, B)\, P(D \mid A, B, C)\, P(E \mid A, B, C, D)$

$P(\text{fever})\, P(\text{cough} \mid \text{fever})$ . . . . .

$P(\text{coronavirus} \mid \ldots)$

Build Joint Distribution Using Chain Rule

Two tools to construct joint distribution

1. Product rule

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

2. Chain rule

$$P(X_1, X_2, \dots, X_n\,) = \prod_i P(\, X_i \mid X_1, \dots, X_{i-1}\,)$$

$P(A, B, C) = P(A)P(\, B \mid A \,)P(\, C \mid A, B\,)$ for ordering A, B, C

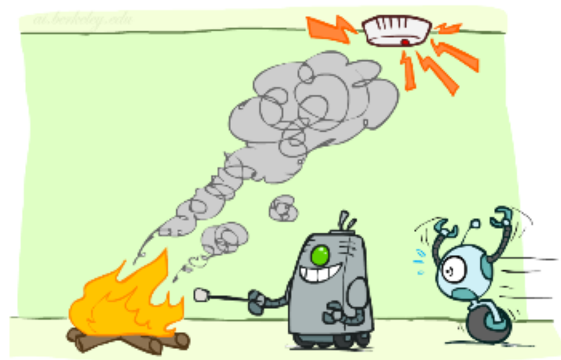
$P(A, B, C) = P(A)P(\, C \mid A \,)P(\, B \mid A, C\,)$ for ordering A, C, B

$P(A, B, C) = P(C)P(\, B \mid C \,)P(\, A \mid C, B\,)$ for ordering C, B, A

Build Joint Distribution Using Chain Rule

Binary random variables

- Fire
- Smoke
- Alarm

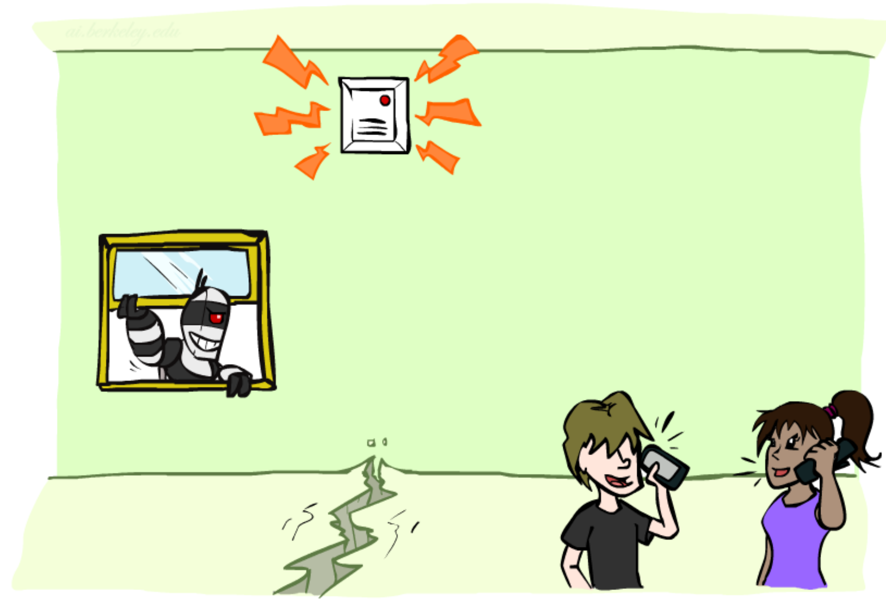


$$P(F,S,A) = P(F)\,P(S|F)\,P(A|S,F)$$

Piazza Poll 3

Variables

- B: Burglary
- A: Alarm goes off
- M: Mary calls
- J: John calls
- E: Earthquake!



How many different ways can we write the chain rule?

A. 1

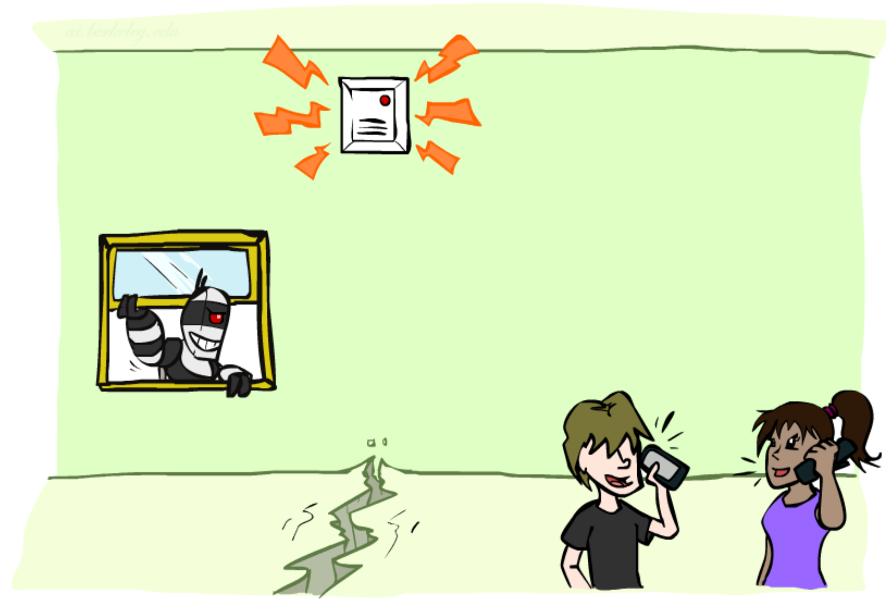
~~*B.* 5~~ Calamity

C. 5 *choose* 5

D. 5!

E. 5^5

Piazza Poll 3



Variables

- B: Burglary
- A: Alarm goes off
- M: Mary calls
- J: John calls
- E: Earthquake!

How many different ways can we write the chain rule?

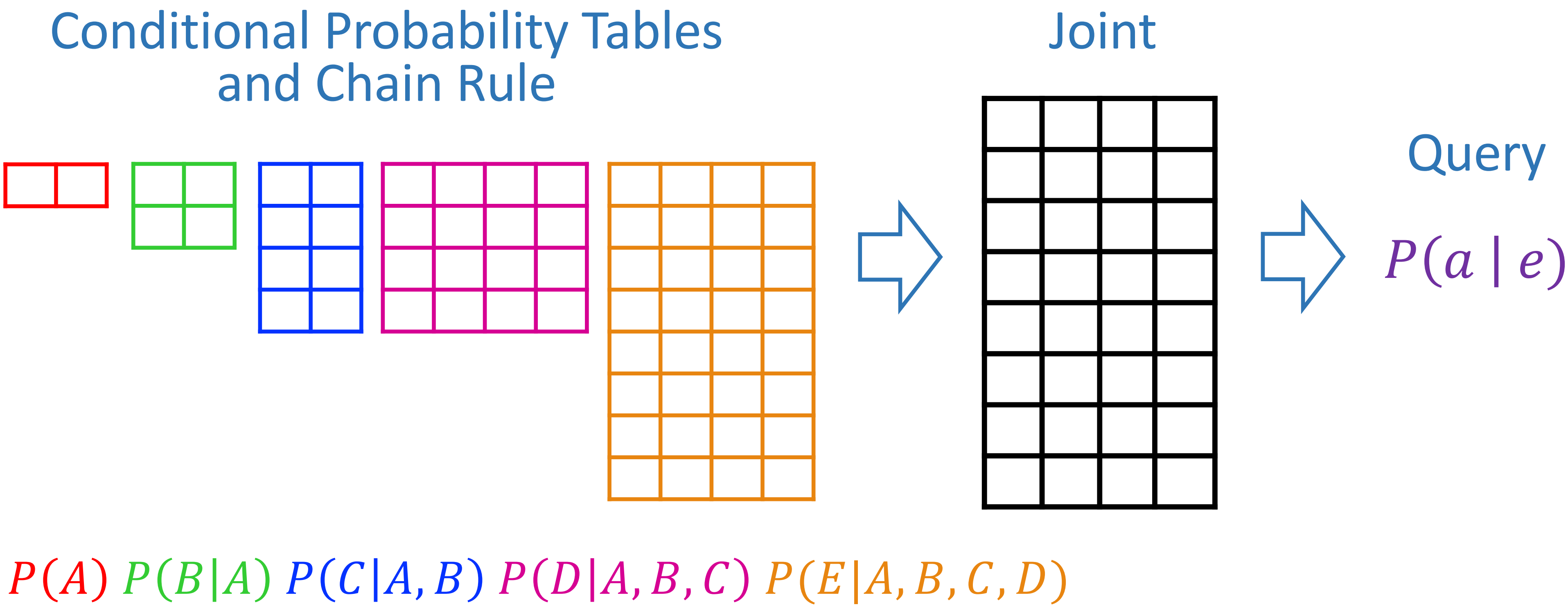
A. 1
B. 5
C. 5 *choose* 5
***D.* 5!**
E. 5^5

$P(B)\,P(A|B)$

B
A
M
J
E

$5 \times 4 \times 3 \times 2 \times 1$

Build Joint Distribution Using Chain Rule



Answer Any Query from Condition Probability Tables

Process to go from (specific) conditional probability tables to query

1. Construct the joint distribution
    1. Product Rule or Chain Rule
2. Answer query from joint
    1. Definition of conditional probability
    2. Law of total probability (marginalization, summing out)

Answer Any Query from Condition Probability Tables

Bayes' rule as an example

Given: $P(E|Q),\ P(Q)$ Query: $P(Q \mid e)$

1. Construct the joint distribution
   1. Product Rule or Chain Rule
$$P(E,Q) = P(E|Q)P(Q)$$
2. Answer query from joint
   1. Definition of conditional probability
$$P(\,Q \mid e\,) = \frac{P(e,Q)}{P(e)}$$
   2. Law of total probability (marginalization, summing out)
$$P(\,Q \mid e\,) = \frac{P(e,Q)}{\sum_q P(e,q)}$$

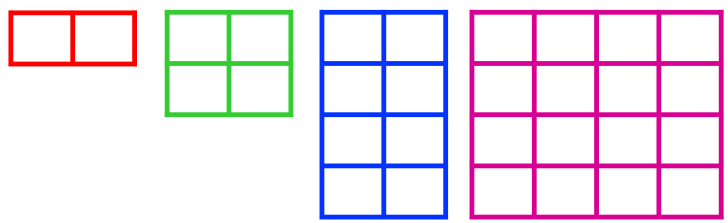
Bayesian Networks

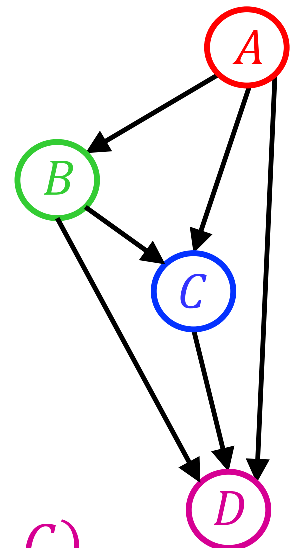
Bayes net

One node per random variable

DAG

One CPT per node: P(node | *Parents*(node))





$$P(A, B, C, D) = P(A)\, P(B|A)\, P(C|A, B)\, P(D|A, B, C)$$

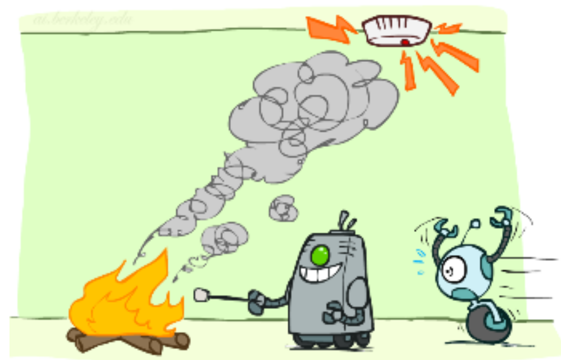
Encode joint distributions as product of conditional distributions on each variable

$$P(X_1, \dots, X_N) = \prod_i P(X_i |\, Parents(X_i))$$

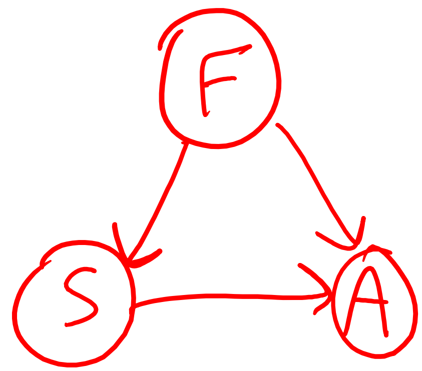
Build Bayes Net Using Chain Rule

Binary random variables

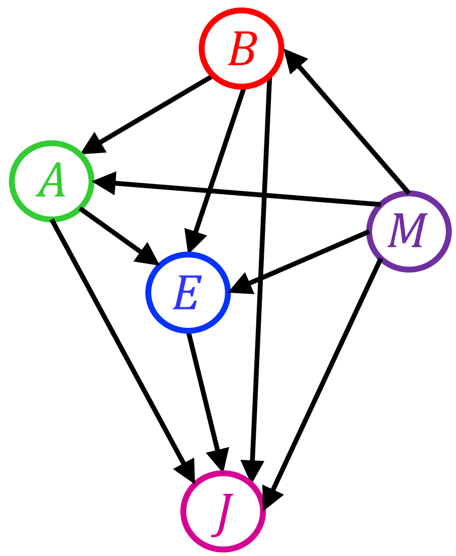
- Fire
- Smoke
- Alarm

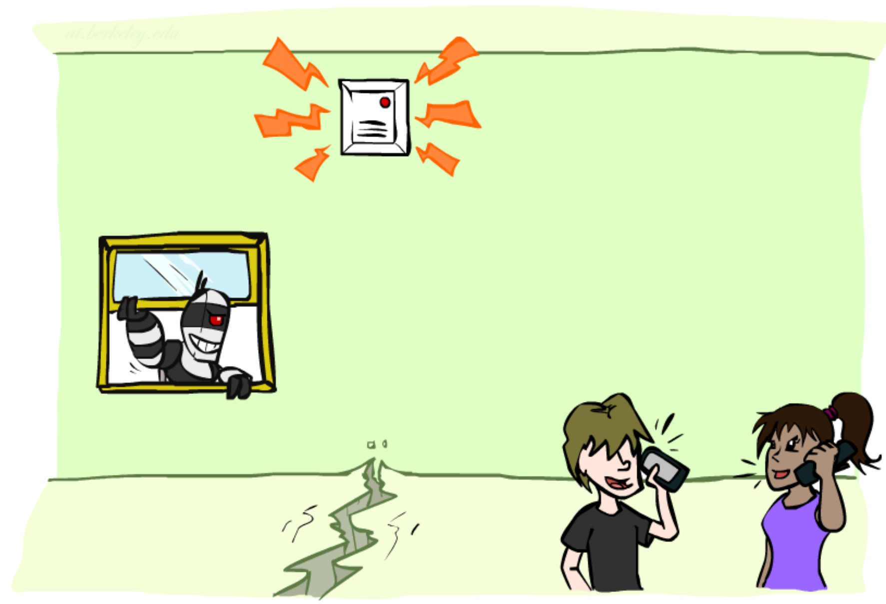


$$P(F, S, A) = P(F)\,P(S|F)\,P(A|F,S)$$



Question





Variables

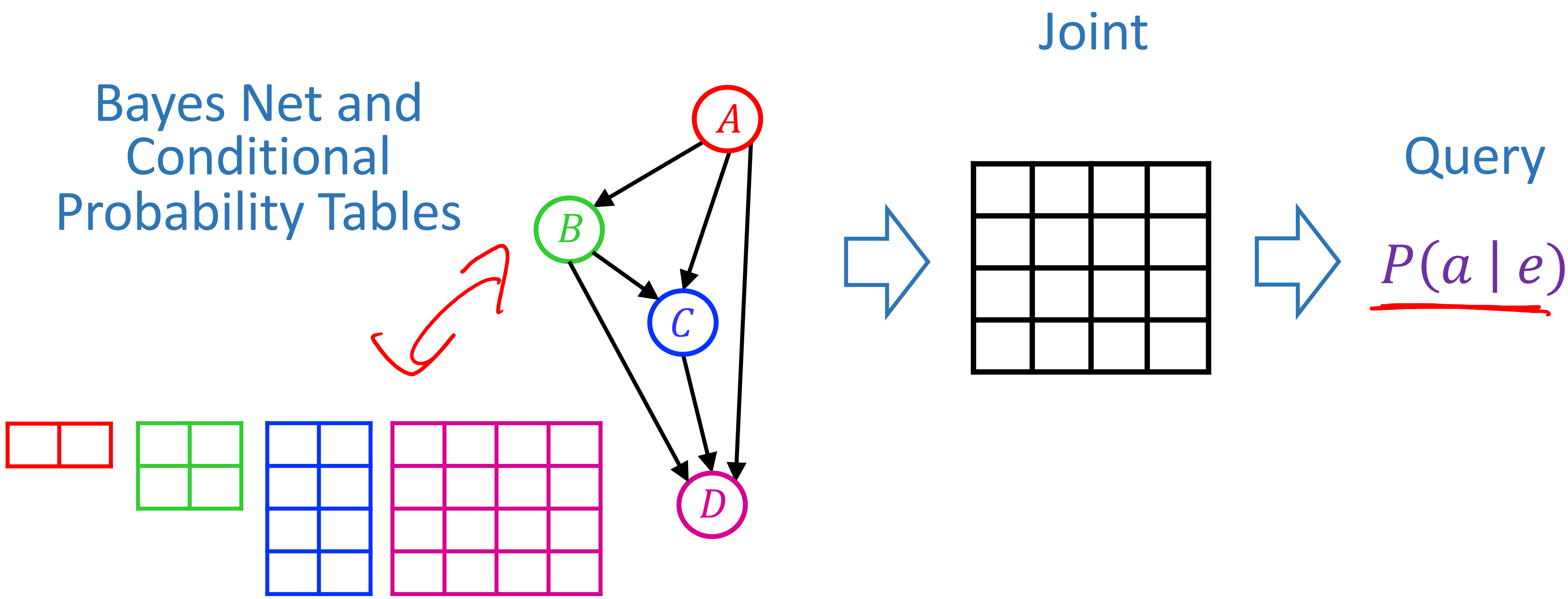
- B: Burglary
- A: Alarm goes off
- M: Mary calls
- J: John calls
- E: Earthquake!

Given the Bayes net, write the joint distribution?

$P(B, A, M, E, J) =$

$P(M)\,P(B \mid M)\,P(A \mid M, B)\,P(E \mid M, B, A)\,P(J \mid B, A, M, E)$

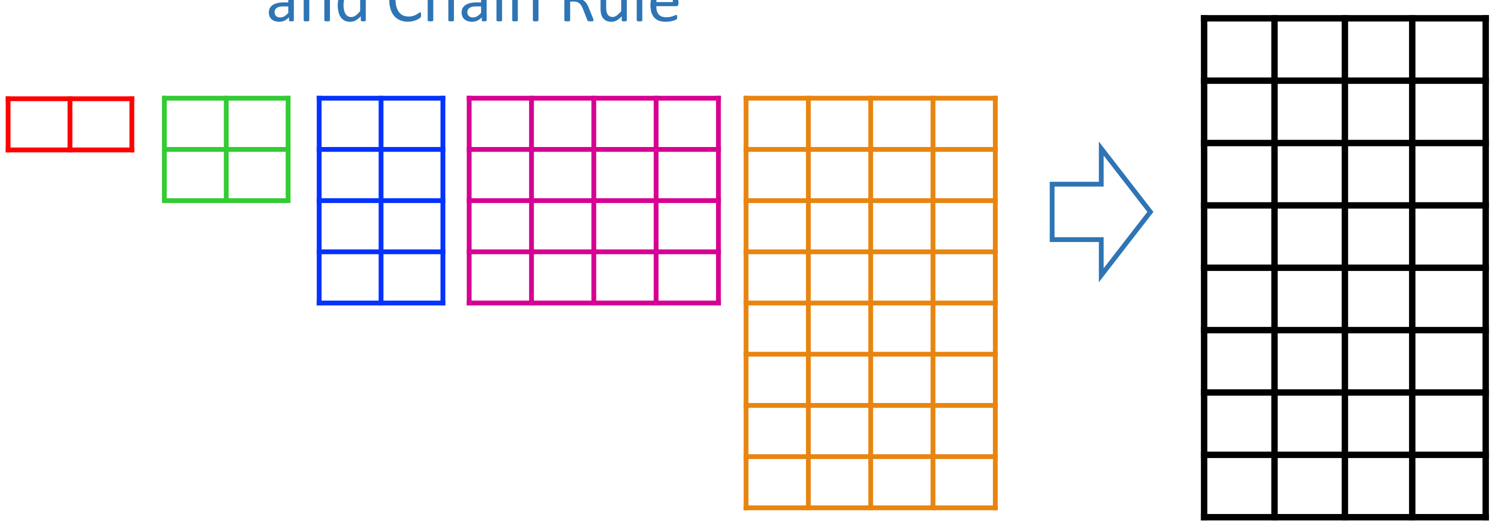
Answer Any Query from Bayes Net



Answer Any Query from Condition Probability Tables

Conditional Probability Tables
and Chain Rule

Joint



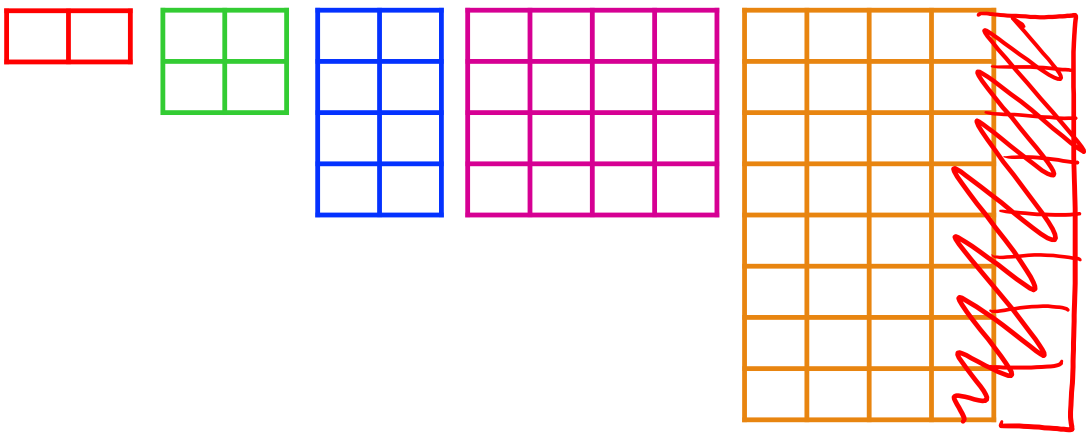
Query

$P(a \mid e)$

$P(A)\ P(B|A)\ P(C|A,B)\ P(D|A,B,C)\ P(E|A,B,C,D)$

Answer Any Query from Condition Probability Tables

Conditional Probability Tables and Chain Rule



$P(A)\ P(B|A)\ P(C|A,B)\ P(D|A,B,C)\ P(E|A,B,C,D)$

Problems

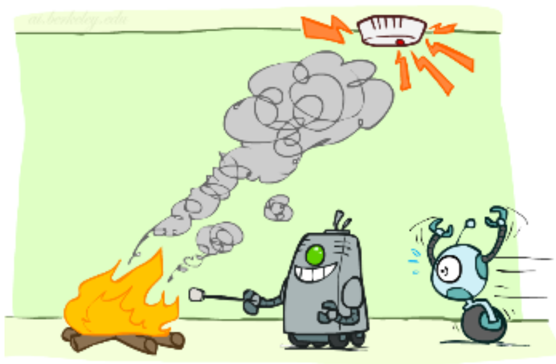
- Huge
  - n variables with d values
  - d^n entries
- We aren't given the right tables

Danielle Belgrave, Microsoft Research

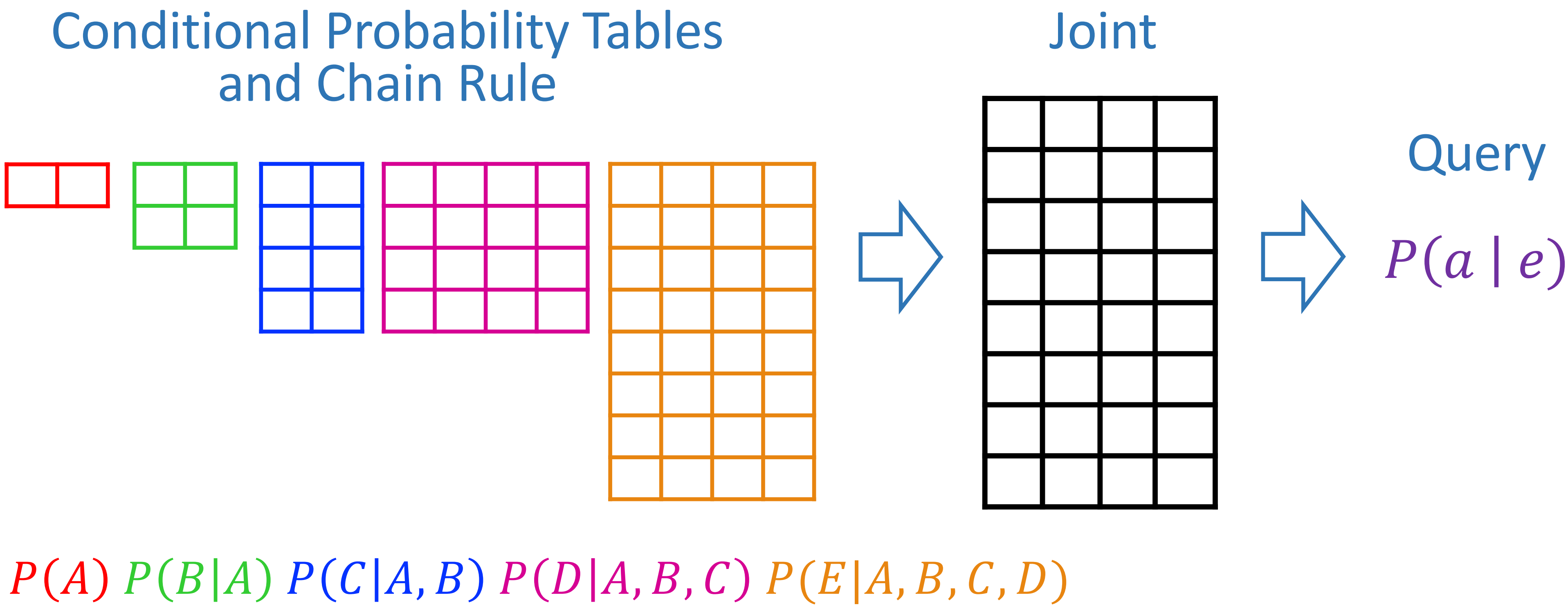
Do We Need the Full Chain Rule?

Binary random variables

- Fire
- Smoke
- Alarm

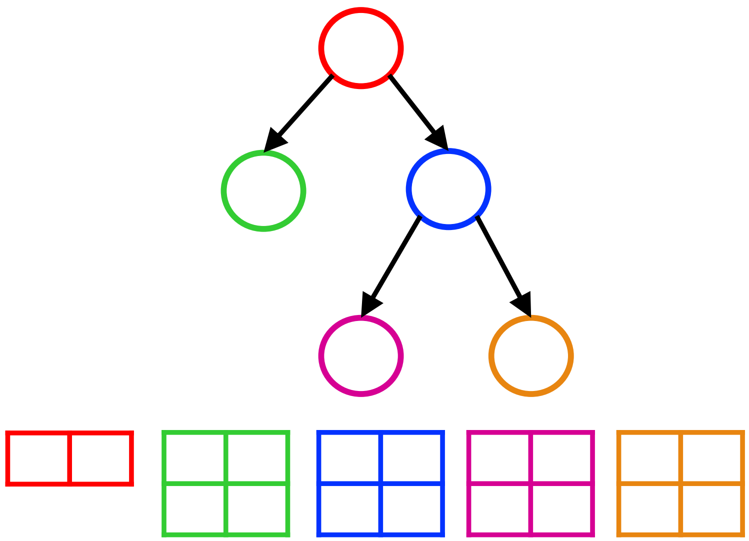


Answer Any Query from Condition Probability Tables



Answer Any Query from Condition Probability Tables

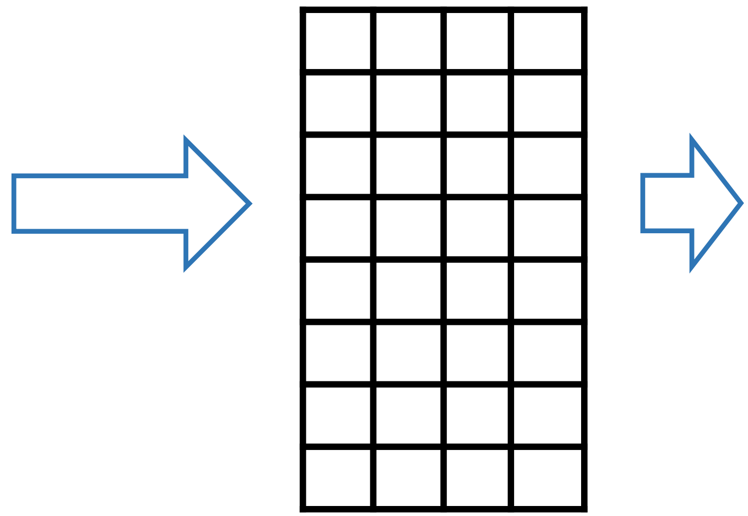
Bayes Net



$P(A)$ $P(B|A)$ $P(C|A)$ $P(D|C)$ $P(E|C)$

$$P(X_1, \dots, X_N) = \prod_i P(X_i | \; Parents(X_i))$$

Joint



Query

$P(a \mid e)$