AI2002 – ARTIFICIAL INTELLIGENCE SPRING 2024 - LECTURE 28-30

Presented By: Mr. Sandesh Kumar

Slides are taken from UC Berkeley

Uncertainty

- Agents may need to handle **uncertainty**, whether due to
  - ***partial observability***,
  - ***nondeterminism***, or
  - ***a combination of the two***.
- An agent may never know for certain what state it’s in or where it will end up after a sequence of actions.
- The right thing to do (**the rational decision**) depends on both the *relative importance of various goals* and *the likelihood that they will be achieved*.
- The agent’s knowledge can at best provide only a **degree of belief** in the relevant sentences.
- Our main tool for dealing with ***degrees of belief*** is **probability theory**.

Probability Basics

- **Logical assertions** say which possible worlds are strictly ruled out.
- **Probabilistic assertions** talk about how probable the various worlds are.

- In probability theory, the set of all possible worlds is called the **sample space** Ω.
- The **possible worlds** are ***mutually exclusive*** and ***exhaustive***
  - two possible worlds cannot both be the case, and one possible world must be the case.
- For example, if we roll two dice, there are 36 possible worlds to consider: (1,1), (1,2), . . ., (6,6).

- A **probability model** associates a numerical probability P(ω) with each possible World ω.
- The basic axioms of probability theory say that every possible world has a probability between 0 and 1 and that the total probability of the set of possible worlds is 1:

$$0 \le P(\omega) \le 1 \text{ for every } \omega \text{ and } \sum_{\omega \in \Omega} P(\omega) = 1$$

Probability Basics

- ***Probabilistic assertions*** are *not usually about particular possible worlds*, but ***about sets of possible worlds.***
- In probability theory, the ***sets of possible words*** are called **events.**
- In AI, these sets are described by **propositions.**

- The **probability associated with a proposition** is defined to be the sum of the probabilities of the worlds in which it holds:

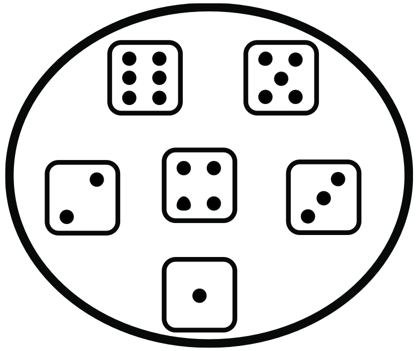
$$\text{For any proposition } \phi,\ P(\phi) = \sum_{\omega \in \phi} P(\omega)$$

- For example, when rolling fair dice, we have

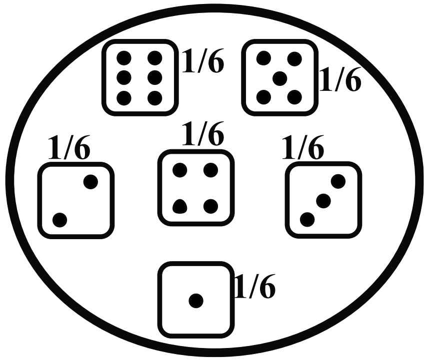
P(Total=11) = P((5, 6)) + P((6, 5)) = 1/36 + 1/36 = 1/18.

Basic laws of probability (discrete)

- Begin with a set Ω of possible worlds
  - E.g., 6 possible rolls of a die, {1, 2, 3, 4, 5, 6}



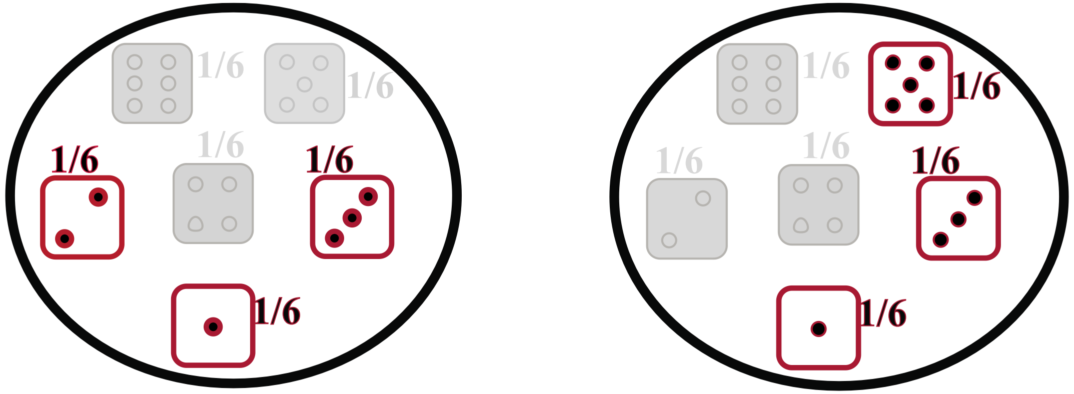
- A ***probability model*** assigns a number $P(\omega)$ to each world ω
  - E.g., $P(1) = P(2) = P(3) = P(5) = P(5) = P(6) = 1/6$.



- These numbers must satisfy
  - $0 \leq P(\omega)$
  - $\sum_{\omega \in \Omega} P(\omega) = 1$

Basic laws contd.

- An ***event*** is any subset of Ω
  - E.g., “roll < 4” is the set {1,2,3}
  - E.g., “roll is odd” is the set {1,3,5}



- The probability of an event is the ***sum*** of probabilities over its worlds
  - $P(A) = \sum_{\omega \in A} P(\omega)$
  - E.g., $P(roll < 4) = P(1) + P(2) + P(3) = 1/2$

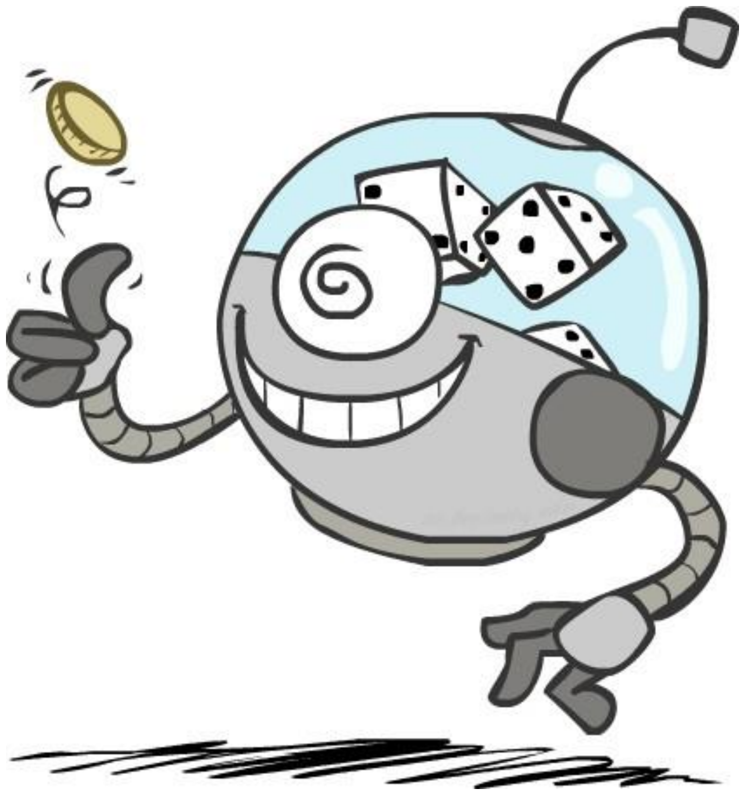
- De Finetti (1931): anyone who bets according to probabilities that violate these laws can be forced to lose money on every set of bets

Random Variables

- A random variable is some aspect of the world about which we (may) have uncertainty
  - R = Is it raining?
  - T = Is it hot or cold?
  - D = How long will it take to drive to work?
  - L = Where is the ghost?

- We denote random variables with capital letters

- Like variables in a CSP, random variables have domains
  - R in {true, false} (often write as {+r, -r})
  - T in {hot, cold}
  - D in $[0, \infty)$
  - L in possible locations, maybe {(0,0), (0,1), …}



Probability Distributions

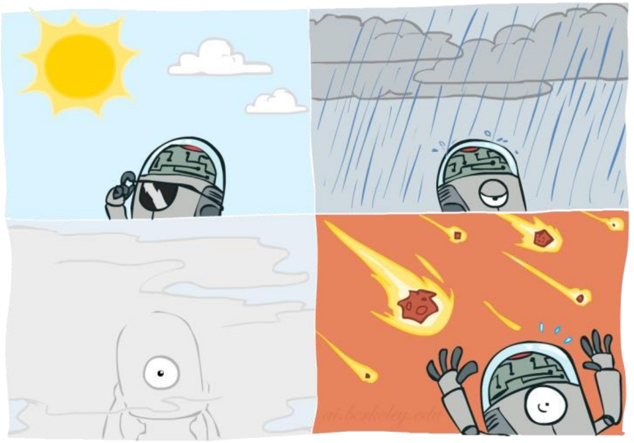
- Associate a probability with each value of that random variable

  - Temperature:

$P(T)$

T	P
hot	0.5
cold	0.5

  - Weather:



$P(W)$

W	P
sun	0.6
rain	0.1
fog	0.3
meteor	0.0

Probability Distributions

- Unobserved random variables have distributions

$P(T)$

T	P
hot	0.5
cold	0.5

$P(W)$

W	P
sun	0.6
rain	0.1
fog	0.3
meteor	0.0

Shorthand notation:

$P(hot)$ same as $P(T = hot)$

$P(cold)$ same as $P(T = cold)$

$P(rain)$ same as $P(W = rain)$

...

OK *if* all domain entries are unique

- A distribution is a TABLE of probabilities of values

- A probability (of a lower case value) is a single number:

$$P(W = rain) = 0.1$$

- Must have: $\forall x \;\; P(X = x) \geq 0$ and $\sum_x P(X = x) = 1$

Joint Distributions

- A *joint distribution* over a set of random variables: $X_1, X_2, \dots, X_N$ specifies a real number for each assignment (or *outcome*):

$$P(X_1 = x_1, X_2 = x_2, \dots, X_N = x_N)$$
$$P(x_1, x_2, \dots, x_N)$$

  - Must obey: $$P(x_1, x_2, \dots x_n) \geq 0$$
$$\sum_{(x_1, x_2, \dots x_n)} P(x_1, x_2, \dots x_n) = 1$$

$P(T, W)$

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

- Size of distribution if n variables with domain sizes d?
  - For all but the smallest distributions, impractical to write out!

Probabilistic Models

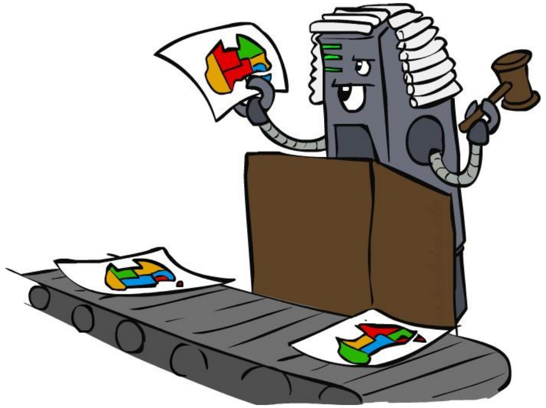
- A probabilistic model is a joint distribution over a set of random variables
- Probabilistic models:
  - (Random) variables with domains
  - Assignments are called *outcomes*
  - Joint distributions: say whether assignments (outcomes) are likely
  - *Normalized:* sum to 1.0
  - Ideally: only certain variables directly interact
- Constraint satisfaction problems:
  - Variables with domains
  - Constraints: state whether assignments are possible
  - Ideally: only certain variables directly interact

Distribution over T,W

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

Constraint over T,W

T	W	P
hot	sun	T
hot	rain	F
cold	sun	F
cold	rain	T



Events

- An *event* is a set E of outcomes

$$P(E) = \sum_{(x_1 \ldots x_n) \in E} P(x_1 \ldots x_n)$$

- From a joint distribution, we can calculate the probability of any event
  - Probability that it's hot AND sunny?
  - Probability that it's hot?
  - Probability that it's hot OR sunny?
- Typically, the events we care about are *partial assignments*, like P(T=hot)

$P(T, W)$

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

Quiz: Events

- P(+x, +y) ?
- P(+x) ?
- P(-y OR +x) ?

$$P(X, Y)$$

X	Y	P
+x	+y	0.2
+x	-y	0.3
-x	+y	0.4
-x	-y	0.1

Quiz: Events

- P(+x, +y) ?   0.2

- P(+x) ?   0.2 + 0.3 = 0.5

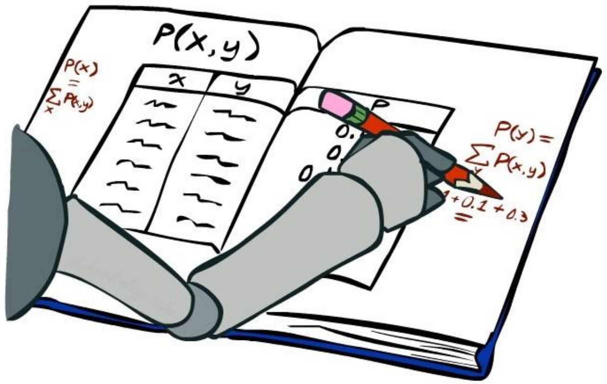
- P(-y OR +x) ?   0.1 + 0.3 + 0.2 = 0.6

$$P(X, Y)$$

X	Y	P
+x	+y	0.2
+x	-y	0.3
-x	+y	0.4
-x	-y	0.1

Marginal Distributions

- Marginal distributions are sub-tables which eliminate random variables
- Marginalization (summing out): Combine collapsed rows by adding



$P(T, W)$

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

$P(t) = 7 \quad P(t, w)$

$\xrightarrow{\$}$

$P(T)$

T	P
hot	0.5
cold	0.5

$P(w) = 7 \quad P(t, w)$

$\xrightarrow{t}$

$P(W)$

W	P
sun	0.6
rain	0.4

$P(X_1 = x_1) = 7 \quad P(X_1 = x_1, X_2 = x_2)$

x_2

hidden (unobserved) variables

Quiz: Marginal Distributions

$P(X, Y)$

X	Y	P
+x	+y	0.2
+x	-y	0.3
-x	+y	0.4
-x	-y	0.1

$P(x) = 7 \quad P(x, y)$

$\xrightarrow{\quad\quad\quad y \quad\quad\quad}$

$P(y) = 7 \quad P(x, y)$

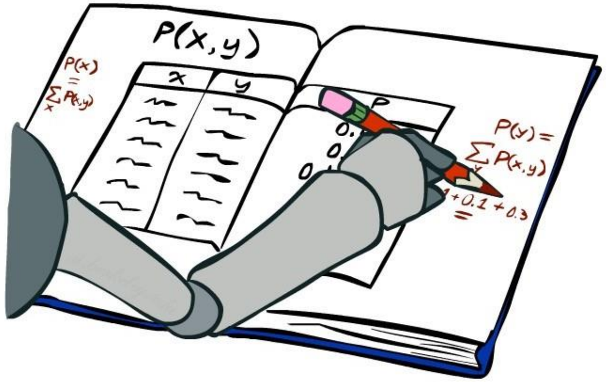
$\xrightarrow{\quad\quad\quad x \quad\quad\quad}$

$P(X)$

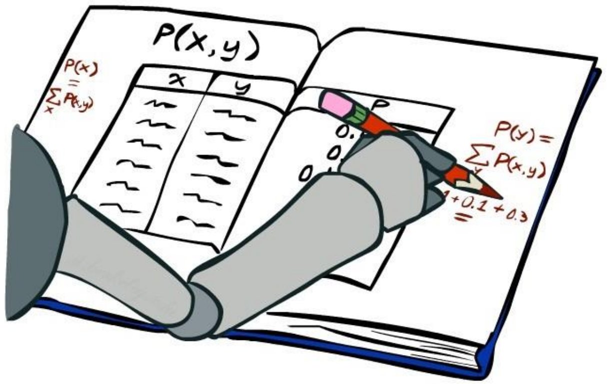
X	P
+x	
-x	

$P(Y)$

Y	P
+y	
-y	



Quiz: Marginal Distributions



$P(X,Y)$

X	Y	P
+x	+y	0.2
+x	-y	0.3
-x	+y	0.4
-x	-y	0.1

$$P(x) = 7 \quad P(x,y)$$

$$\xrightarrow{\quad y \quad}$$

$$P(y) = 7 \quad P(x,y)$$

$$\xrightarrow{\quad x \quad}$$

$P(X)$

X	P
+x	0.5
-x	0.5

$P(Y)$

Y	P
+y	0.6
-y	0.4

Probability Basics:
Unconditional Probability and Conditional Probability

- Probabilities such as P(event) are called **unconditional** or **prior probabilities.**
- **Prior probabilities** refer to **degrees of belief in propositions** in ***the absence of any other information***.

- Most of the time, however, we have some information, usually called **evidence**, that has already been revealed.
- In that case, we are interested not in the unconditional probability of an **event**, but the **conditional** or **posterior probability** of **event** given **evidence**.
- **Conditional probability** is written as **P(event| evidence)**

Conditional Probability

- Conditional probabilities are defined in terms of unconditional probabilities as follows: for any propositions a and b, we have

$$P(a \mid b) = \frac{P(a \wedge b)}{P(b)}$$

which holds whenever P(b) > 0.

- The definition of conditional probability can be written in a different form called the **product rule**:

$$P(a \wedge b) = P(a \mid b)P(b)$$

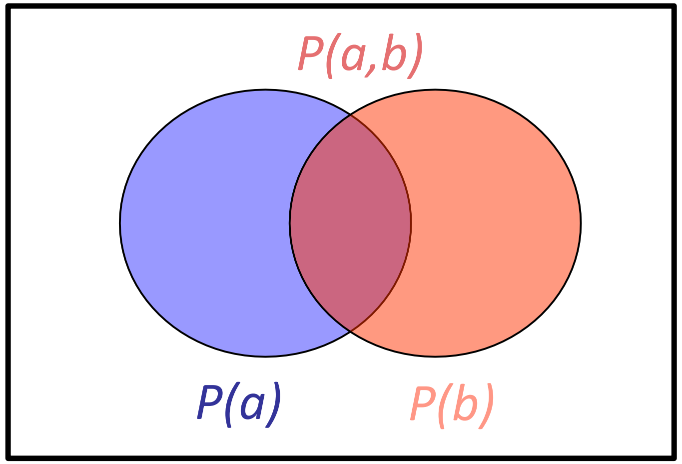
Conditional Probabilities

- A simple relation between joint and conditional probabilities
  - In fact, this is taken as the *definition* of a conditional probability

evidence

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

query = (proportion of *b* where *a* holds)

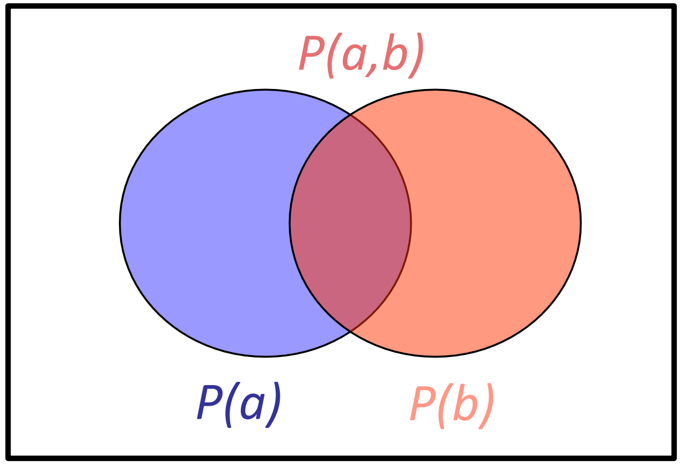


Conditional Probabilities

- A simple relation between joint and conditional probabilities
  - In fact, this is taken as the *definition* of a conditional probability

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

= (proportion of *b* where *a* holds)



$$P(T,W)$$

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

$$P(W=s|T=c) = \frac{P(W=s,T=c)}{P(T=c)} = \frac{0.2}{0.5} = 0.4$$

$$= P(W=s,T=c) + P(W=r,T=c)$$
$$= 0.2 + 0.3 \quad = 0.5$$

Quiz: Conditional Probabilities

$$P(X, Y)$$

X	Y	P
+x	+y	0.2
+x	-y	0.3
-x	+y	0.4
-x	-y	0.1

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

- P(+x | +y) ?
- P(-x | +y) ?
- P(-y | +x) ?

Quiz: Conditional Probabilities

$$P(X, Y)$$

X	Y	P
+x	+y	0.2
+x	-y	0.3
-x	+y	0.4
-x	-y	0.1

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

- P(+x | +y) ? 0.2 / 0.6 = 1/3
- P(-x | +y) ? 0.4 / 0.6 = 2/3
- P(-y | +x) ? 0.3 / 0.5 = 3/5

Conditional Distributions

- Conditional distributions are probability distributions over some variables given fixed values of others

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

Conditional Distributions

$P(W|T)$

$$P(W|T = hot)$$

W	P
sun	0.8
rain	0.2

$$P(W|T = cold)$$

W	P
sun	0.4
rain	0.6

Joint Distribution

$$P(T, W)$$

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

To Normalize

- (Dictionary) To bring or restore to a normal condition

All entries sum to ONE

- Procedure:
  - Step 1: Compute Z = sum over all entries
  - Step 2: Divide every entry by Z

- Example 1

W	P
sun	0.2
rain	0.3

Normalize

Z = 0.5

W	P
sun	0.4
rain	0.6

- Example 2

T	W	P
hot	sun	20
hot	rain	5
cold	sun	10
cold	rain	15

Normalize

Z = 50

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

Normalization Trick

$$P(T, W)$$

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

$$\begin{aligned} P(W=s|T=c) &= \frac{P(W=s,T=c)}{P(T=c)} \\ &= \frac{P(W=s,T=c)}{P(W=s,T=c)+P(W=r,T=c)} \\ &= \frac{0.2}{0.2+0.3} = 0.4 \end{aligned}$$

$$\begin{aligned} P(W=r|T=c) &= \frac{P(W=r,T=c)}{P(T=c)} \\ &= \frac{P(W=r,T=c)}{P(W=s,T=c)+P(W=r,T=c)} \\ &= \frac{0.3}{0.2+0.3} = 0.6 \end{aligned}$$

$$P(W|T=c)$$

W	P
sun	0.4
rain	0.6

Normalization Trick

$$
\begin{aligned}
P(W=s|T=c) &= \frac{P(W=s,T=c)}{P(T=c)} \\
&= \frac{P(W=s,T=c)}{P(W=s,T=c)+P(W=r,T=c)} \\
&= \frac{0.2}{0.2+0.3} = 0.4
\end{aligned}
$$

$P(T,W)$

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

SELECT the joint probabilities matching the evidence

$P(c,W)$

T	W	P
cold	sun	0.2
cold	rain	0.3

NORMALIZE the selection (make it sum to one)

$P(W|T=c)$

W	P
sun	0.4
rain	0.6

$$
\begin{aligned}
P(W=r|T=c) &= \frac{P(W=r,T=c)}{P(T=c)} \\
&= \frac{P(W=r,T=c)}{P(W=s,T=c)+P(W=r,T=c)} \\
&= \frac{0.3}{0.2+0.3} = 0.6
\end{aligned}
$$

Normalization Trick

$P(T, W)$

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

SELECT the joint probabilities matching the evidence

$P(c, W)$

T	W	P
cold	sun	0.2
cold	rain	0.3

NORMALIZE the selection (make it sum to one)

$P(W|T = c)$

W	P
sun	0.4
rain	0.6

- Why does this work? Sum of selection is P(evidence)! (P(T=c), here)

$$P(x_1|x_2) = \frac{P(x_1, x_2)}{P(x_2)} = \frac{P(x_1, x_2)}{\sum_{x_1} P(x_1, x_2)}$$

Quiz: Normalization Trick

- P(X | Y=-y) ?

$$P(X, Y)$$

X	Y	P
+x	+y	0.2
+x	-y	0.3
-x	+y	0.4
-x	-y	0.1

SELECT the joint probabilities matching the evidence

NORMALIZE the selection (make it sum to one)

Quiz: Normalization Trick

- P(X | Y=-y) ?

$P(X, Y)$

X	Y	P
+x	+y	0.2
+x	-y	0.3
-x	+y	0.4
-x	-y	0.1

SELECT the joint probabilities matching the evidence

X	Y	P
+x	-y	0.3
-x	-y	0.1

NORMALIZE the selection (make it sum to one)

X	P
+x	0.75
-x	0.25

Probabilistic Inference

- *Probabilistic inference:* compute a desired probability from other known probabilities (e.g. conditional from joint)

- We generally compute conditional probabilities
  - P(on time | no reported accidents) = 0.90
  - These represent the agent's *beliefs* given the evidence

- Probabilities change with new evidence:
  - P(on time | no accidents, 5 a.m.) = 0.95
  - P(on time | no accidents, 5 a.m., raining) = 0.80
  - Observing new evidence causes *beliefs to be updated*

Inference by Enumeration

- P(W)?

S	T	W	P
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

Inference by Enumeration

- P(W)?

  query

S	T	W	P
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

Inference by Enumeration

- P(W)?

$$P(sun) = 0.3 + 0.1 + 0.1 + 0.15 = 0.65$$

S	T	W	P
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

Inference by Enumeration

- P(W)?

P(sun) = 0.3 + 0.1 + 0.1 + 0.15 = 0.65

P(rain) = 0.05 + 0.05 + 0.05 + 0.20 = 0.35

S	T	W	P
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

Inference by Enumeration

evidence

- P(W | winter, hot)?

query

S	T	W	P
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

Inference by Enumeration

- P(W | winter, hot)?

unnormalized P(sun | winter, hot) = 0.10

unnormalized P(rain | winter, hot) = 0.05

P(sun | winter, hot) = 0.10 / 0.15 = 2/3

P(rain | winter, hot) = 0.05 / 0.15 = 1/3

S	T	W	P
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

Inference by Enumeration

evidence

- P(W | winter)?

query

hidden (unobserved) variable: T

S	T	W	P
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

Inference by Enumeration

- P(W | winter)?

  unnormalized P(sun | winter) = 0.1 + 0.15 = 0.25

S	T	W	P
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

Inference by Enumeration

- P(W | winter)?

  unnormalized P(sun | winter) = 0.1 + 0.15 = 0.25

  unnormalized P(rain | winter) = 0.05 + 0.20 = 0.25

S	T	W	P
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

Inference by Enumeration

- P(W | winter)?

  unnormalized P(sun | winter) = 0.1 + 0.15 = 0.25

  unnormalized P(rain | winter) = 0.05 + 0.20 = 0.25

  P(sun | winter) = 0.25 / 0.50 = 0.5

  P(rain | winter) = 0.25 / 0.50 = 0.5

S	T	W	P
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

Inference by Enumeration

- General case:
  - Evidence variables: $E_1 \dots E_k = e_1 \dots e_k$
  - Query* variable: Q
  - Hidden variables: $H_1 \dots H_r$

$$\left.\begin{array}{l} E_1 \dots E_k = e_1 \dots e_k \\ Q \\ H_1 \dots H_r \end{array}\right\} X_1, X_2, \dots X_n$$

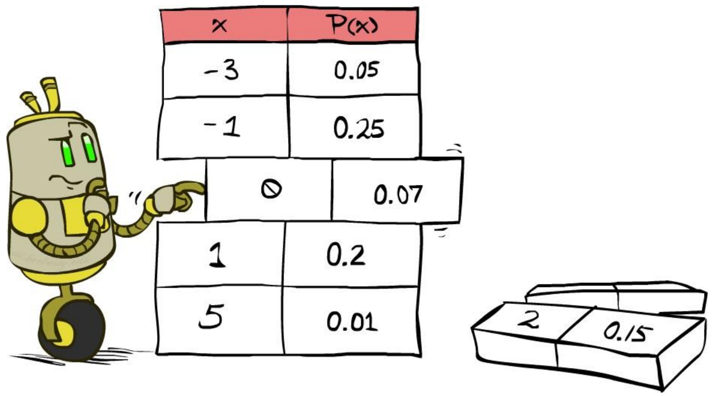
All variables

- We want:

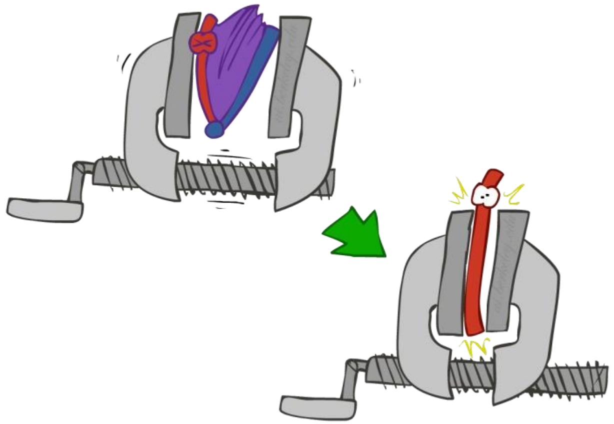
** Works fine with multiple query variables, too*

$$P(Q|e_1 \dots e_k)$$

- Step 1: **Select** the entries consistent with the evidence



- Step 2: **Sum** out H to get joint of Query and evidence



$$P(Q, e_1 \dots e_k) = \sum_{h_1 \dots h_r} \underbrace{P(Q, h_1 \dots h_r, e_1 \dots e_k)}_{X_1, X_2, \dots X_n}$$

- Step 3: **Normalize**

$$\times \frac{1}{Z}$$

$$Z = \sum_q P(Q, e_1 \cdots e_k)$$

$$P(Q|e_1 \cdots e_k) = \frac{1}{Z} P(Q, e_1 \cdots e_k)$$

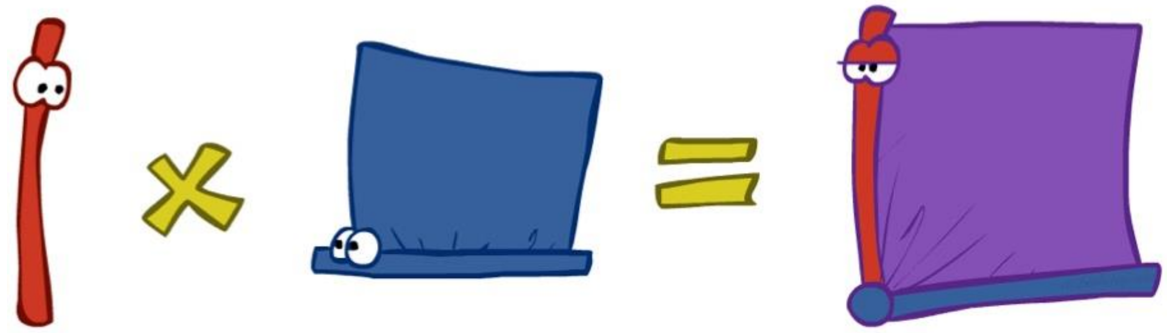
Inference by Enumeration

- Obvious problems:
  - Worst-case time complexity $O(d^n)$
  - Space complexity $O(d^n)$ to store the joint distribution

The Product Rule

- Sometimes have conditional distributions but want the joint

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



The Product Rule

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

- Example:

$P(W)$

R	P
sun	0.8
rain	0.2

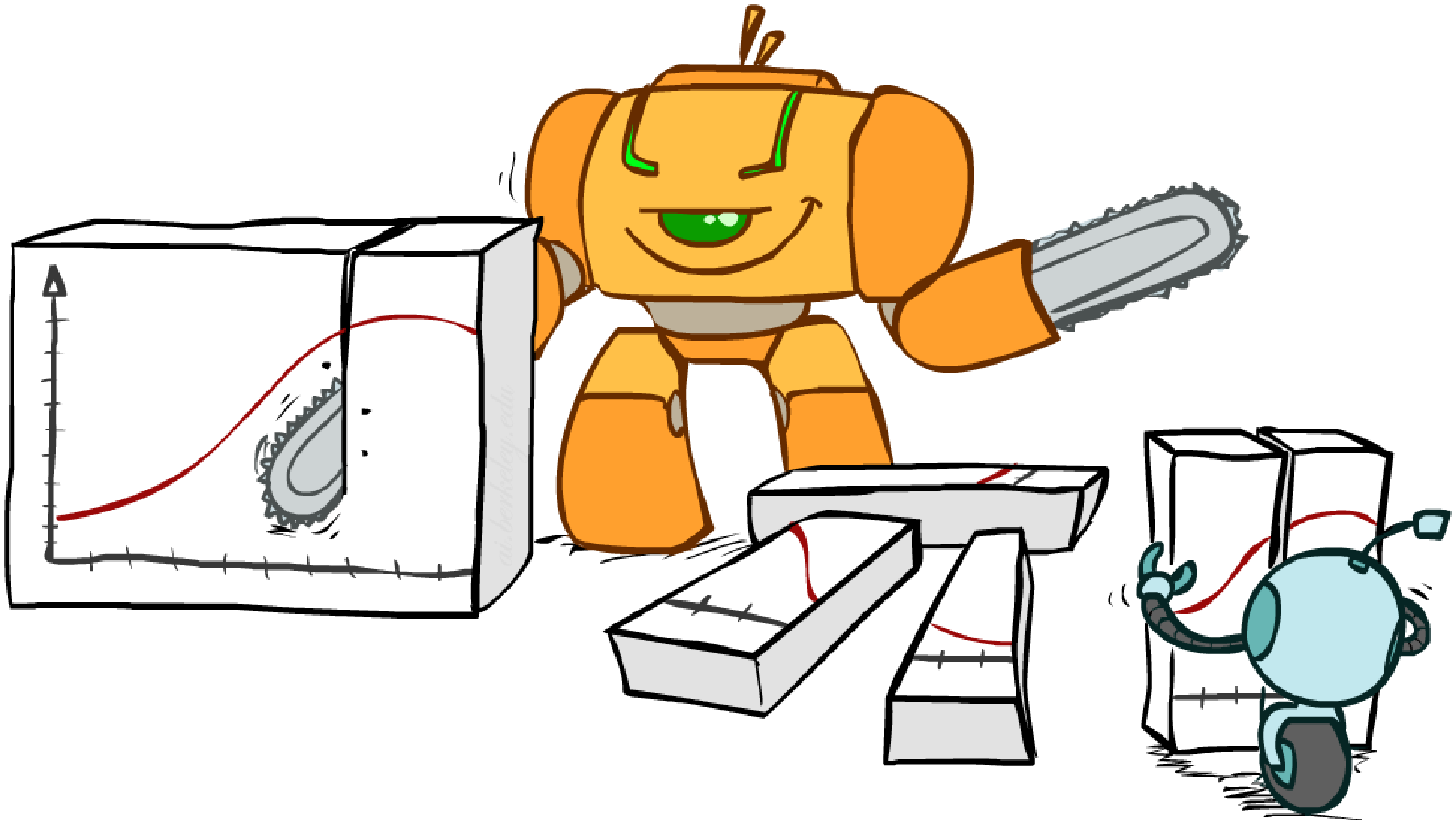
$P(D|W)$

D	W	P
wet	sun	0.1
dry	sun	0.9
wet	rain	0.7
dry	rain	0.3

$P(D,W)$

D	W	P
wet	sun	
dry	sun	
wet	rain	
dry	rain	

Bayes' Rule



Bayes' Rule

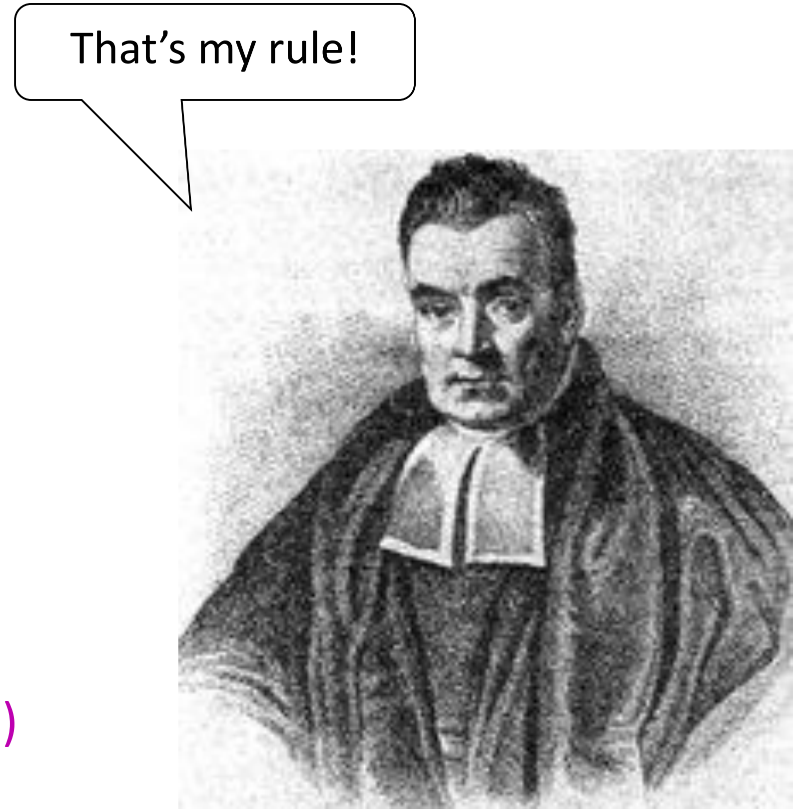
- Write the product rule both ways:

$$P(a \mid b)\, P(b) = P(a, b) = P(b \mid a)\, P(a)$$

- Dividing left and right expressions, we get:

$$P(a \mid b) = \frac{P(b \mid a)\, P(a)}{P(b)}$$

- Why is this at all helpful?
  - Lets us build one conditional from its reverse
  - Often one conditional is tricky but the other one is simple
  - Describes an "update" step from prior $P(a)$ to posterior $P(a \mid b)$
    - Hence provides a simple, formal theory of learning



- E.g.: d is data, h is hypothesis: $P(h \mid d) = \frac{P(d \mid h)\, P(h)}{P(d)} = \frac{P(d \mid h)\, P(h)}{\Sigma_{h'} P(d|h')\, P(h')}$

Inference with Bayes' Rule

- Example: Diagnostic probability from causal probability:

$$P(cause \mid effect) = \frac{P(effect \mid cause)\, P(cause)}{P(effect)}$$

- Example:
  - M: meningitis, S: stiff neck

$$\left.\begin{array}{l} P(s \mid m) = 0.8 \\ P(m) = 0.0001 \\ P(s) = 0.01 \end{array}\right\} \text{Example givens}$$

$$P(m \mid s) = \frac{P(s \mid m)\, P(m)}{P(s)} = \frac{0.8 \times 0.0001}{0.01}$$

  - Note: posterior probability of meningitis still very small: 0.008 (80x bigger – why?)
  - Note: you should still get stiff necks checked out! Why?

Quiz: Bayes' Rule

- Given:

$P(W)$

R	P
sun	0.8
rain	0.2

$P(D|W)$

D	W	P
wet	sun	0.1
dry	sun	0.9
wet	rain	0.7
dry	rain	0.3

- What is P(W | dry) ?

Quiz: Bayes' Rule

- Given:

$P(W)$

R	P
sun	0.8
rain	0.2

$P(D|W)$

D	W	P
wet	sun	0.1
dry	sun	0.9
wet	rain	0.7
dry	rain	0.3

- What is P(W | dry) ?

unnormalized P(sun|dry) = P(dry|sun) * P(sun) = 0.9 * 0.8 = 0.72

unnormalized P(rain|dry) = P(dry|rain) * P(rain) = 0.3 * 0.2 = 0.06

P(sun|dry)= 0.72/0.78 = 12/13

P(rain|dry)= 0.06/0.78 = 1/13