Uncertainty

- observed variables (evidence) : agents know certain things about state of the world
- unobserved variables : agent has to reason
- model : how observed and unobserved variables relate

Random variables

aspect of world of which we may have uncertainty

notation:

- capital letters
- R in {true, false} often written as {+r, -r}
- T in {hot, cold}

domains

Probability distribution

associate probability with each value

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

distribution

not the best

shorthand notation

$P(hot) = P(T = hot)$

ok if all domain entries are unique

probability is a single number $P(W = rain) = 0.1$

Must have $\forall x \in P(X = x) \geq 0 \ \wedge \ \sum P(X = x) = 1$

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.

$P(X_1 = x_1 \dots X_n = x_n)$

$P(x_1, x_2 \dots x_n)$

all positive and sum to 1

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

- size of distribution n variables with domain size d? d^n

we don't build it entirely

Probabilistic models

a joint distribution over a set of random var.

- (Random) variables with domains
- Assignments called outcomes
- joint distributions say if outcomes are likely
- normalized (add to 1)
- ideally only certain variables interact

Events

a set E of outcomes

$$P(E) = \sum_{(x_1 \dots x_n) \in E} P(x_1, x_2, x_n)$$

from a joint distribution we can calculate the probability of any event

Marginal distributions

sub tables that eliminate variables

marginalization: combine collapsed rows by adding

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

$P(t) = \sum_s P(t,s)$ $\Longrightarrow$

P(t)	P
hot	0.5
cold	0.5

$P(s) = \sum_t P(t,s)$ $\Longrightarrow$

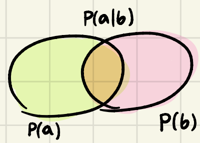
P(w)

W	P
sun	0.6
rain	0.4

Conditional probabilities

simple relation between joint and conditional probabilities

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



P(T, W)

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

$$P(W = sun \mid T = cold) = \frac{0.2}{0.5} = \frac{0.2}{0.2 + 0.3}$$

↖ important for later

Conditional distributions

probability distributions over some variables given fixed values of others

joint distribution

$P(T,W)$

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

conditional distribution

$P(W|T=hot)$

W	P
sun	0.8
rain	0.2

$P(W|T=cold)$

W	P
sun	0.4
rain	0.2

Normalization trick

$P(T,W)$

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

$$P(W=s|T=c) = \frac{P(W=s,T=c)}{P(T=c)} = \frac{P(W=s,T=c)}{P(W=s,T=c)+P(W=r,T=c)} = \frac{0.2}{0.2+0.3}$$

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

$$P(W=r|T=c) = \frac{P(W=r,T=c)}{P(T=c)} = \frac{P(W=r,T=c)}{P(W=s,T=c)+P(W=r,T=c)}$$

same denominator and it is the sum of the two numerators

Algorithm $P(W|T=c)$

1) select rows that are consistent with the evidence

$P(T,W)$

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

adds up to 1

$\Longrightarrow$

$P(c,W)$

T	W	P
cold	sun	0.2
cold	rain	0.3

$\Downarrow$ divide by sum (0.5)

2) normalize

$P(W|T=c)$

W	P
sun	0.4
rain	0.6

Probabilistic inference

compute a probability from others that are known

usually we compute conditional probabilities

probability changes with new evidence

Inference by enumeration

- general case:
  - evidence var. $E_1 \dots E_k = e_1 \dots e_k$
  - query var. Q
  - hidden var. $H_1 \dots H_r$

$\left.\right\}$ $X_1, X_2 \dots X_n$ All variables

we want: $P(Q \mid e_1 \dots e_k)$

1) select entries consistent with the evidence
2) Sum out H to get joint of query and evidence
3) Normalize $\times \frac{1}{Z}$

problems: time $O(d^n)$
space $O(d^n)$

The product rule

have conditional distribution but want the

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

P(W)

W	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

$\iff$

P(D,W)

D	W	P
wet	sun	0.08
dry	sun	0.72
wet	rain	0.14
dry	rain	0.06

$\leftarrow 0.1 \cdot 0.8$

The chain rule

we can write any joint distribution as an incremental product of conditional distributions

$$P(x_1, x_2, x_3) = P(x_1)\,P(x_2|x_1)\,P(x_3|x_1,x_2)$$

$$P(x_1 \dots x_n) = \prod P(x_i \mid x_1 \dots x_{i-1})$$

Bayes rule

two ways to factor a joint distribution over two variables:

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

diving we get:

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