CSE 473

Chapter 13

More Uncertainty

Outline for Next Few Lectures

- Basic notions
  - Atomic events, probabilities, joint distribution
  - Inference by enumeration
  - Independence & conditional independence
  - Bayes' rule
- Bayesian networks
- Statistical learning

Logic vs. Probability

Symbol: Q, R …	**Random variable:** Q …
Boolean values: T, F	**Values/Domain:** you specify e.g. {heads, tails} [1,6]
State of the world: Assignment of T/F to all Q, R … Z	**Atomic event:** a complete assignment of values to Q… Z • Mutually exclusive • Exhaustive
	Prior probability (aka Unconditional prob: P(Q)
	Joint distribution: Prob. of every atomic event

 3

Types of Random Variables

Propositional or Boolean random variables

e.g., $Cavity$ (do I have a cavity?)

Discrete random variables (*finite* or *infinite*)

e.g., $Weather$ is one of $\langle sunny, rain, cloudy, snow \rangle$

$Weather = rain$ is a proposition

Values must be exhaustive and mutually exclusive

Continuous random variables (*bounded* or *unbounded*)

e.g., $Temp = 21.6$; also allow, e.g., $Temp < 22.0$.

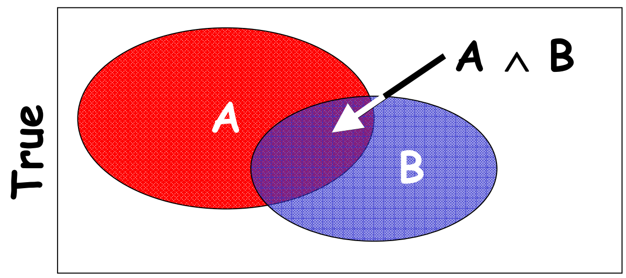
Arbitrary Boolean combinations of basic propositions

 4

Axioms of Probability Theory

- Just 3 are enough to build entire theory!
  1. All probabilities between 0 and 1
     $0 \le P(A) \le 1$
  2. P(true) = 1 and P(false) = 0
  3. Probability of disjunction of events is:

$$P(A \vee B) = P(A) + P(B) - P(A \wedge B)$$



Prior and Joint Probability

Prior or unconditional probabilities of propositions
e.g., $P(Cavity = true) = 0.2$ and $P(Weather = sunny) = 0.72$
correspond to belief prior to arrival of any (new) evidence

Probability distribution gives values for all possible assignments:
$\mathbf{P}(Weather) = \langle 0.72, 0.1, 0.08, 0.1 \rangle$ (*normalized*, i.e., sums to 1)

Joint probability distribution for a set of r.v.s gives the probability of every atomic event on those r.v.s
$\mathbf{P}(Weather, Cavity) =$ a 4×2 matrix of values:

$Weather =$	$sunny$	$rain$	$cloudy$	$snow$
$Cavity = true$	0.144	0.02	0.016	0.02
$Cavity = false$	0.576	0.08	0.064	0.08

We will see later how any question can be answered by the joint distribution

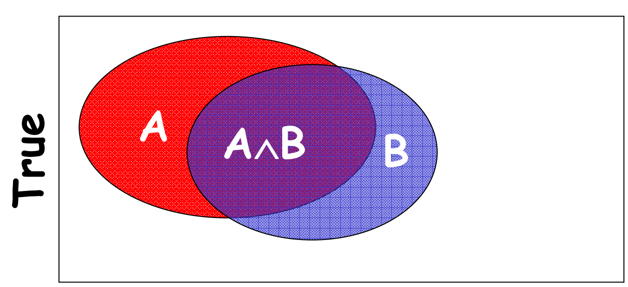
Conditional (or Posterior) Probability

- Conditional or posterior probabilities
  e.g., P(*cavity* | *toothache*) = 0.8
  i.e., given that *Toothache* is true (and all I know)
- Notation for conditional **distributions**:
  P(*Cavity* | *Toothache*) = 2-element vector of 2-element vectors (2 P values when Toothache is true and 2 when false)
- If we know more, e.g., *cavity* is also given, then we have
  P(*cavity* | *toothache, cavity*) = 1
- New evidence may be irrelevant, allowing simplification:
  P(*cavity* | *toothache, sunny*) = P(*cavity* | *toothache*) = 0.8

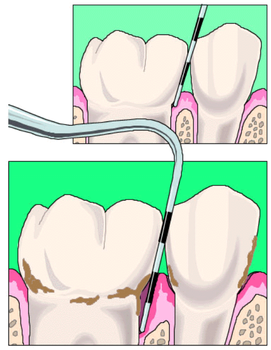
Conditional Probability

- P(*A* | *B*) is the probability of *A* given *B*
- Assumes that *B* is the only info known.
- Defined as:

$$P(A \mid B) = \frac{P(A \wedge B)}{P(B)}$$



Dilemma at the Dentist's



What is the probability of a cavity given a toothache?
What is the probability of a cavity given the probe catches?

Probabilistic Inference by Enumeration

Start with the joint distribution:

	toothache		¬ *toothache*	
	catch	¬ *catch*	*catch*	¬ *catch*
cavity	.108	.012	.072	.008
¬ *cavity*	.016	.064	.144	.576

For any proposition ϕ, sum the atomic events where it is true:

$$P(\phi) = \Sigma_{\omega:\omega \models \phi} P(\omega)$$

P(toothache)= .108+.012+.016+.064
= .20 or 20%

Inference by Enumeration

Start with the joint distribution:

	toothache		¬ *toothache*	
	catch	¬ *catch*	*catch*	¬ *catch*
cavity	**.108**	**.012**	**.072**	**.008**
¬ *cavity*	**.016**	**.064**	**.144**	**.576**

P(toothache∨cavity) = .20 + .108 + .012 + .072 + .008-(.108+.012)

= .28

Inference by Enumeration

Start with the joint distribution:

	toothache		¬ *toothache*	
	catch	¬ *catch*	*catch*	¬ *catch*
cavity	**.108**	**.012**	**.072**	**.008**
¬ *cavity*	**.016**	**.064**	**.144**	**.576**

Can also compute conditional probabilities:

$$P(\neg cavity | toothache) = \frac{P(\neg cavity \wedge toothache)}{P(toothache)}$$

$$= \frac{0.016 + 0.064}{0.108 + 0.012 + 0.016 + 0.064} = 0.4$$

Problems with Enumeration

- Worst case time: $O(d^n)$
  Where d = max arity of random variables
  e.g., d = 2 for Boolean (T/F)
  And n = number of random variables
- Space complexity also $O(d^n)$
  Size of joint distribution
- Problem: Hard/impossible to estimate all $O(d^n)$ entries for large problems

Independence

- *A* and *B* are *independent* iff:

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

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

These two constraints are logically equivalent

- Therefore, if *A* and *B* are independent:

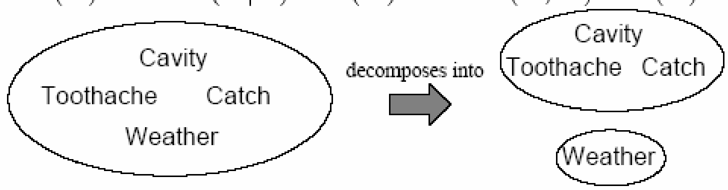
$$P(A \mid B) = \frac{P(A \wedge B)}{P(B)} = P(A)$$

$$P(A \wedge B) = P(A)P(B)$$

Independence

A and B are independent iff

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



$\mathbf{P}(Toothache, Catch, Cavity, Weather)$

$= \mathbf{P}(Toothache, Catch, Cavity)\mathbf{P}(Weather)$

2 values → 4 values

32 entries reduced to 12; for n independent biased coins, $2^n \to n$

Complete independence is powerful but rare

What to do if it doesn't hold?

 15

Conditional Independence

$\mathbf{P}(Toothache, Cavity, Catch)$ has $2^3 - 1 = 7$ independent entries

If I have a cavity, the probability that the probe catches in it doesn't depend on whether I have a toothache:

(1) $P(catch|toothache, cavity) = P(catch|cavity)$

The same independence holds if I haven't got a cavity:

(2) $P(catch|toothache, \neg cavity) = P(catch|\neg cavity)$

$Catch$ is *conditionally independent* of $Toothache$ given $Cavity$:

$\mathbf{P}(Catch|Toothache, Cavity) = \mathbf{P}(Catch|Cavity)$

Instead of 7 entries, only need 5 (why?)

 16

Conditional Independence II

P(catch | toothache, cavity) = P(catch | cavity)
P(catch | toothache,¬cavity) = P(catch |¬cavity)

Equivalent statements:

$$\mathbf{P}(Toothache|Catch, Cavity) = \mathbf{P}(Toothache|Cavity)$$
$$\mathbf{P}(Toothache, Catch|Cavity) = \mathbf{P}(Toothache|Cavity)\mathbf{P}(Catch|Cavity)$$

Why only 5 entries in table?

Write out full joint distribution using chain rule:

$$\begin{aligned}&\mathbf{P}(Toothache, Catch, Cavity)\\ &= \mathbf{P}(Toothache|Catch, Cavity)\mathbf{P}(Catch, Cavity)\\ &= \mathbf{P}(Toothache|Catch, Cavity)\mathbf{P}(Catch|Cavity)\mathbf{P}(Cavity)\\ &= \mathbf{P}(Toothache|Cavity)\mathbf{P}(Catch|Cavity)\mathbf{P}(Cavity)\end{aligned}$$

I.e., $2 + 2 + 1 = 5$ independent numbers

Power of Cond. Independence

- Often, using conditional independence reduces the storage complexity of the joint distribution from exponential to linear!!
- Conditional independence is the most basic & robust form of knowledge about uncertain environments.

Next Time

- Bayes' Rule
- Bayesian Inference
- Bayesian Networks