Introduction to Artificial Intelligence

Uncertainty

Chapter 14

Dieter Fox

Outline

◇ Bayes' rule

◇ Independence

◇ Robot localization

Bayes' Rule

Product rule $P(A \land B) = P(A|B)P(B) = P(B|A)P(A)$

$\Rightarrow$ **Bayes' rule** $P(A|B) = \dfrac{P(B|A)P(A)}{P(B)}$

Why is this useful???

For assessing **diagnostic** probability from **causal** probability:

$$P(Cause|Effect) = \frac{P(Effect|Cause)P(Cause)}{P(Effect)}$$

E.g., let M be meningitis, S be stiff neck:

$$P(M|S) = \frac{P(S|M)P(M)}{P(S)} = \frac{0.8 \times 0.0001}{0.1} = 0.0008$$

Note: posterior probability of meningitis still very small!

Normalization

Suppose we wish to compute a posterior distribution over A given $B = b$, and suppose A has possible values $a_1 \dots a_m$

We can apply Bayes' rule for each value of A:

$P(A = a_1|B = b) = P(B = b|A = a_1)P(A = a_1)/P(B = b)$

$\dots$

$P(A = a_m|B = b) = P(B = b|A = a_m)P(A = a_m)/P(B = b)$

Adding these up, and noting that $\sum_i P(A = a_i|B = b) = 1$:

$$1/P(B = b) = 1/\sum_i P(B = b|A = a_i)P(A = a_i)$$

This is the **normalization factor**, constant w.r.t. i, denoted α:

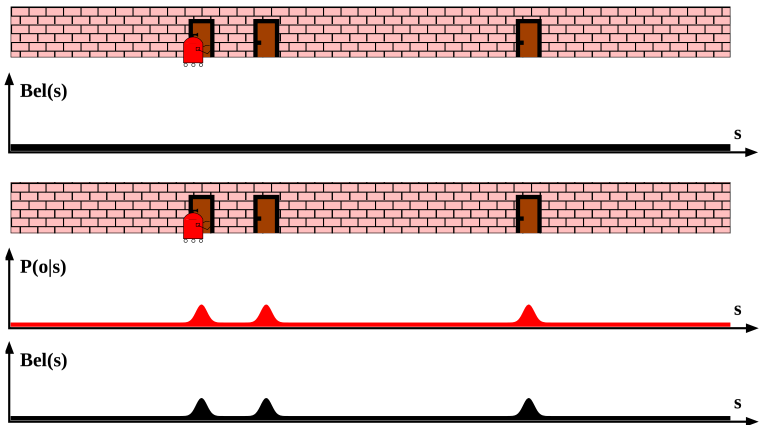
$$\mathbf{P}(A|B = b) = \alpha \mathbf{P}(B = b|A)\mathbf{P}(A)$$

Typically compute an unnormalized distribution, normalize at end

e.g., suppose $\mathbf{P}(B = b|A)\mathbf{P}(A) = \langle 0.4, 0.2, 0.2 \rangle$

then $\mathbf{P}(A|B = b) = \alpha \langle 0.4, 0.2, 0.2 \rangle = \frac{\langle 0.4, 0.2, 0.2 \rangle}{0.4+0.2+0.2} = \langle 0.5, 0.25, 0.25 \rangle$

Application of Bayes' Rule



$$P(s|o) = \frac{P(o|s)P(s)}{P(o)} = \frac{P(o|s)P(s)}{\sum_s P(o|s)P(s)}$$

Full joint distributions

A **complete probability model** specifies every entry in the joint distribution for all the variables $\mathbf{X} = X_1, \ldots, X_n$

I.e., a probability for each possible world $X_1 = x_1, \ldots, X_n = x_n$

E.g., suppose $Toothache$ and $Cavity$ are the random variables:

	$Toothache = true$	$Toothache = false$
$Cavity = true$	0.04	0.06
$Cavity = false$	0.01	0.89

Possible worlds are mutually exclusive $\Rightarrow$ $P(w_1 \wedge w_2) = 0$

Possible worlds are exhaustive $\Rightarrow$ $w_1 \vee \cdots \vee w_n$ is $True$

hence $\sum_i P(w_i) = 1$

Full joint distributions contd.

1) For any proposition ϕ defined on the random variables $\phi(w_i)$ is true or false

2) ϕ is equivalent to the disjunction of w_is where $\phi(w_i)$ is true

Hence $P(\phi) = \sum_{\{w_i:\ \phi(w_i)\}} P(w_i)$

I.e., the unconditional probability of any proposition is computable as the sum of entries from the full joint distribution

Conditional probabilities can be computed in the same way as a ratio:

$$P(\phi|\xi) = \frac{P(\phi \wedge \xi)}{P(\xi)}$$

E.g.,

$$P(Cavity|Toothache) = \frac{P(Cavity \wedge Toothache)}{P(Toothache)} = \frac{0.04}{0.04 + 0.01} = 0.8$$

Independence

Two random variables A B are (absolutely) **independent** iff

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

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

e.g., A and B are two coin tosses

If n Boolean variables are independent, the full joint is

$\mathbf{P}(X_1, \ldots, X_n) = \prod_i \mathbf{P}(X_i)$

hence can be specified by just n numbers

Absolute independence is a very strong requirement, seldom met

Conditional independence

Consider the dentist problem with three random variables:
$Toothache$, $Cavity$, $Catch$ (steel probe catches in my tooth)

The full joint distribution 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) $P(Catch|Toothache, Cavity) = P(Catch|Cavity)$
i.e., $Catch$ is **conditionally independent** of $Toothache$ given $Cavity$

The same independence holds if I haven't got a cavity:
(2) $P(Catch|Toothache, \neg Cavity) = P(Catch|\neg Cavity)$

Conditional independence contd.

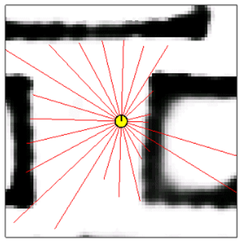
Equivalent statements to
(1) $P(Catch|Toothache, Cavity) = P(Catch|Cavity)$

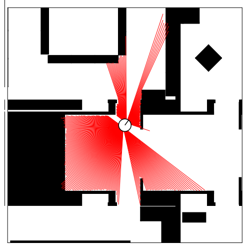
(1a) $P(Toothache|Catch, Cavity) = P(Toothache|Cavity)$

(1b) $P(Toothache, Catch|Cavity) = P(Toothache|Cavity)P(Catch|Cavity)$

Full joint distribution can now be written as
$$\mathbf{P}(Toothache, Catch, Cavity) = \mathbf{P}(Toothache, Catch|Cavity)\mathbf{P}(Cavity)$$
$$= \mathbf{P}(Toothache|Cavity)\mathbf{P}(Catch|Cavity)\mathbf{P}(Cavity)$$
i.e., 2 + 2 + 1 = 5 independent numbers

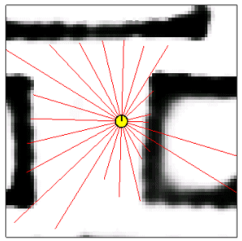
Robot localization with proximity sensors

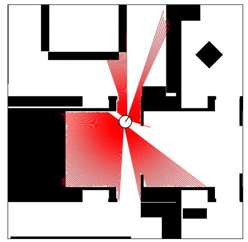




What is $P(s \mid o)$??

Robot localization with proximity sensors

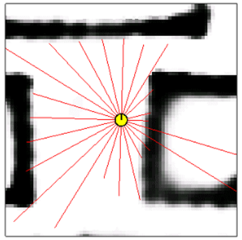


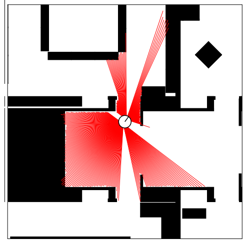


What is $P(s \mid o)$??

$$P(s|o) = \frac{P(o|s)P(s)}{P(o)} = \frac{P(o|s)P(s)}{\sum_s P(o|s)P(s)}$$

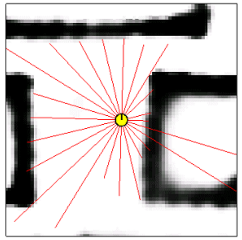
What's the probability of a sensor scan?

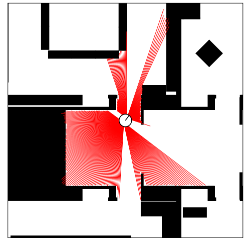




How can we get $P(o \mid s)$??

What's the probability of a sensor scan?

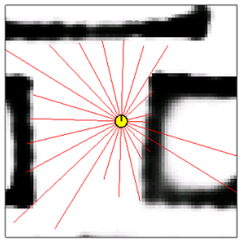


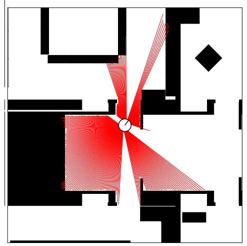


How can we get $P(o \mid s)$??

$P(o \mid s) = P(o_1, o_2, \ldots, o_n \mid s)$

What's the probability of a sensor scan?





How can we get $P(o \mid s)$??

$P(o \mid s) = P(o_1, o_2, \ldots, o_n \mid s)$

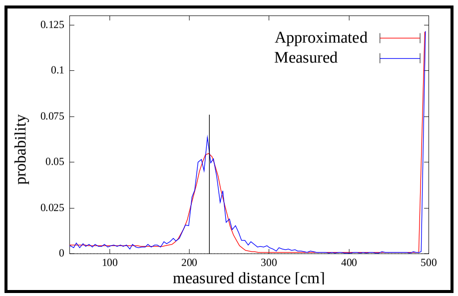
Assumption: sensor beams are conditionally independent **given** the map and the robot's position

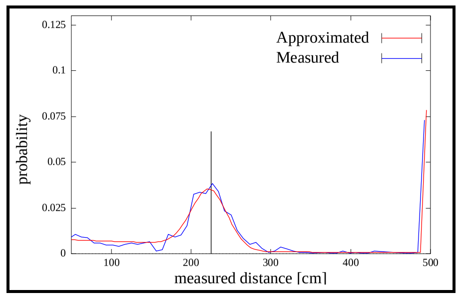
$\Rightarrow \; P(o_1, o_2, \ldots, o_n \mid s) = P(o_1 \mid s) \ldots P(o_n \mid s)$

What's the probability of a sensor beam?

The sensor is either reflected by an **unknown obstacle** or by the **next obstacle in the map**

$$P(d_i \mid s) = 1 - (1 - (1 - \sum_{j<i} P_u(d_j))\, c_d\, P_m(d_i \mid s))) \cdot (1 - (1 - \sum_{j<i} P(d_j))\, c_r)$$





Probability of a laser scan