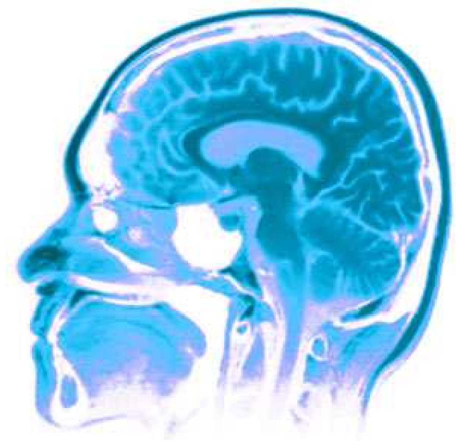
CPSC340



Bayes rule

Nando de Freitas
September, 2012
University of British Columbia

Outline of the lecture

This lecture builds on the concepts of **conditioning** and **marginalization** to introduce **Bayes rule**. The goal is for you to:

- ❑ Learn how Bayes rule is derive.
- ❑ Learn to apply Bayes rule to practical examples.
- ❑ Practice marginalization and conditioning.

Problem 1: Diagnoses

- The doctor has bad news and good news.

- The bad news is that you tested positive for a serious disease, and that **the test is 99% accurate** (i.e., the probability of testing positive given that you have the disease is 0.99, as is the probability of testing negative given that you don't have the disease).

- The good news is that this is a rare disease, striking only 1 in 10,000 people.

- What are the chances that you actually have the disease?

50 5

Problem 2: Monty Hall

On a game show, a contestant is told the rules as follows:

- There are three doors, labeled 1, 2, 3. A single big prize has been hidden behind one of them. The other two doors have goats. You get to select one door.

- Initially your chosen door will not be opened. Instead, the host will open one of the other two doors, and he will do so in such a way as not to reveal the prize.

- At this point, you will be given a fresh choice of door: you can either stick with your first choice, or you can switch to the other closed door. All the doors will then be opened and you will receive whatever is behind your final choice.

Imagine that the contestant chooses door 1 first; then the host opens door 3, revealing a goat behind the door. Should the contestant

(a) stick with door 1, 6
(b) switch to door 2, 50
(c) does it make no difference? 4

Problem 2: Monty Hall

Problem 3: Speech synthesis and recognition

1. *Can computers produce a voice signal given some typed words?*

100%

2. *Can computers produce words when they hear a voice signal?*

3

3. Why?

Bayes rule

Bayes rule enables us to reverse probabilities:

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

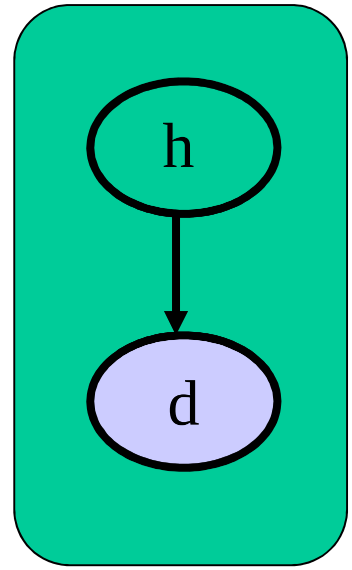
$$P(AB) = P(B|A)\,P(A)$$
$$= P(A|B)\,P(B)$$

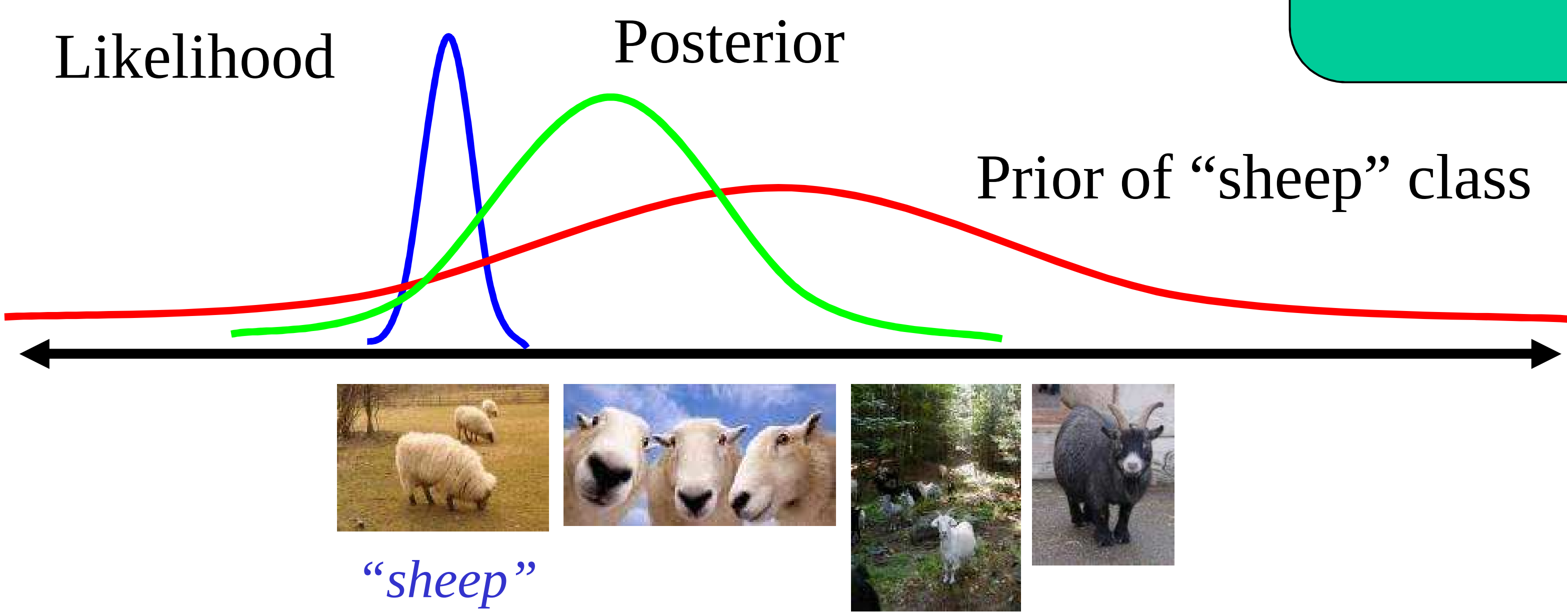
$$P(B|A)P(A) = P(A|B)P(B)$$

$$\sum_B P(B|A) = 1 \qquad \sum_B P(B) = 1$$

Learning and Bayesian inference

$$p(h \mid d) = \frac{p(d \mid h)\, p(h)}{\sum_{h' \in H} p(d \mid h')\, p(h')}$$





Problem 1: Diagnoses

The test is 99% accurate: ***P(T=1|D=1)*** *= 0.99 and* ***P(T=0|D=0)*** *= 0.99*
Where ***T*** *denotes test and* ***D*** *denotes disease.*

The disease affects 1 in 10000: ***P(D=1)*** *= 0.0001*

$$P(D=1|T=1) = \frac{P(T=1|D=1)P(D=1)}{P(T=1|D=0)P(D=0)+P(T=1|D=1)P(D=1)}$$

Josh Tenenbaum

Alison Gopnik

TED

$$\approx 0.0098$$

Problem 2: Monty Hall

(i) ***H=i*** *denote the* ***H****ypothesis that the prize is behind door* ***i****.* ***A priori*** *all 3 doors are equally likely to have the prize:*

$$P(H=1) = P(H=2) = P(H=3) = 1/3$$

(ii) *Contestant* ***chooses door 1.***

Let's think. If the prize is truly behind door 1, the host is indifferent and will choose doors 2 or 3 with equal probability. If the prize is behind door 2 (or 3), host chooses 3 (or 2).

$$P(D=2|H=1) = ½,\ \ P(D=3|H=1) = ½$$
$$P(D=2|H=2) = 0,\ \ P(D=3|H=2) = 1$$
$$P(D=2|H=3) = 1,\ \ P(D=3|H=3) = 0$$

(iii) *The host* ***opens door 3 (D=3)****, revealing a goat behind the door. That is, the observation is* ***D=3. Now is the prize behind door 2 or 1?***

Problem 2: Monty Hall

$P(H=1) = P(H=2) = P(H=3) = 1/3$

$P(D=2|H=1) = ½$, $P(D=3|H=1) = ½$
$P(D=2|H=2) = 0$, $P(D=3|H=2) = 1$
$P(D=2|H=3) = 1$, $P(D=3|H=3) = 0$

We use Bayes rule to compute the probability of the hypothesis that the prize is behind door i (for $i=1,2,3$) given that the host has opened door 3 ($D=3$). That is we compute $P(H=i|D=3)$.

$$P(H=1|D=3) = \frac{P(D=3|H=1)\,P(H=1)}{P(D=3)} = \frac{(1/2)(1/3)}{P(D=3)} = \frac{\frac{1}{6}}{\frac{1}{6}+\frac{1}{3}} = \frac{1}{3}$$

$$P(H=2|D=3) = \frac{P(D=3|H=2)\,P(H=2)}{P(D=3)} = \frac{1\,(1/3)}{P(D=3)} = \frac{1/3}{1/6+1/3} = \frac{2}{3}$$

$$P(H=3|D=3) = \frac{P(D=3|H=3)\,P(H=3)}{P(D=3)} = \frac{0\,(1/3)}{P(D=3)} = 0$$

Problem 2: Monty Hall

Since $P(H=2 \mid D=3) > P(H=1 \mid D=3)$, the contestant should switch.

NOTE: I used the fact that

$$\sum_{i=1}^{3} P(H=i \mid D=3) = P(H=1 \mid D=3) + P(H=2 \mid D=3) + P(H=3 \mid D=3) = 1$$

to <u>normalize</u>. An alternative is to compute

$$P(D=3) = \sum_i P(H=i, D=3) = \sum_i P(D=3 \mid H=i) P(H=i)$$

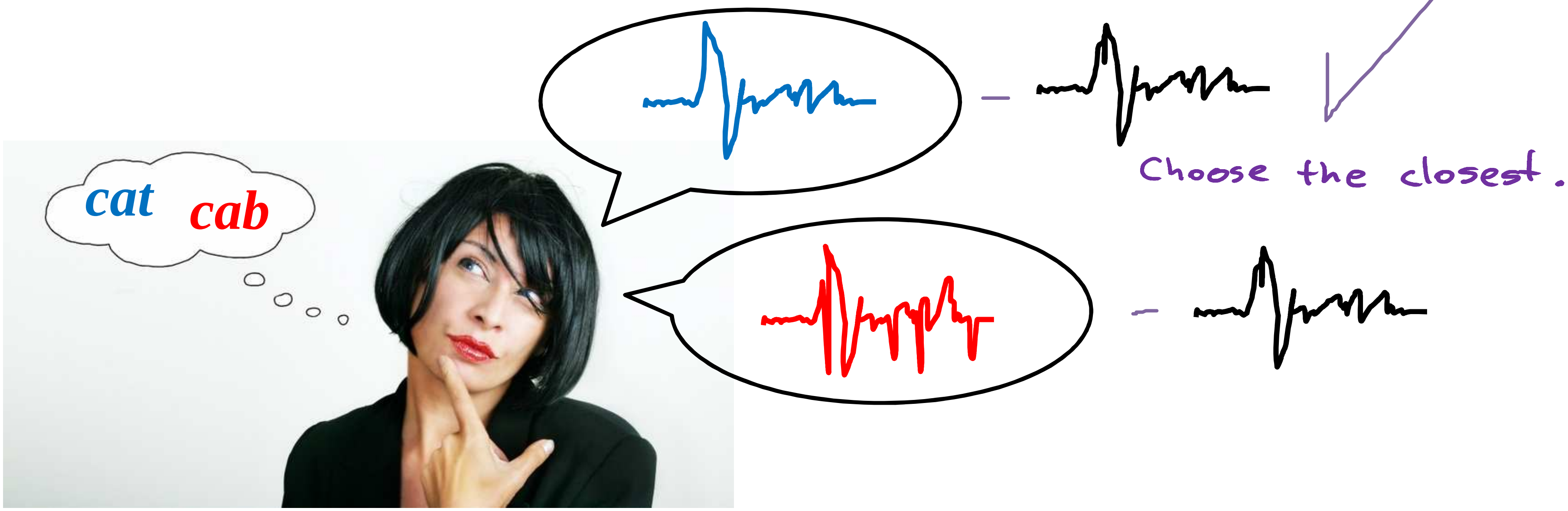
$$= P(D=3 \mid H=1) P(H=1) + P(D=3 \mid H=2) P(H=2)$$

$$+ P(D=3 \mid H=3) P(H=3) = \frac{1}{6} + \frac{1}{3} + 0$$

Problem 3: Speech synthesis and recognition

Suppose she hears:  *Is it cat or cab?*

She imagines the sounds each word produces and compares the sounds



Take Home Message: *Learning takes place to ensure that our **observations** match our **hallucinations***

Speech recognition

$$P(\text{words} \mid \text{sound}) \;\alpha\; P(\text{sound} \mid \text{words})\, P(\text{words})$$

Final beliefs

Likelihood of data

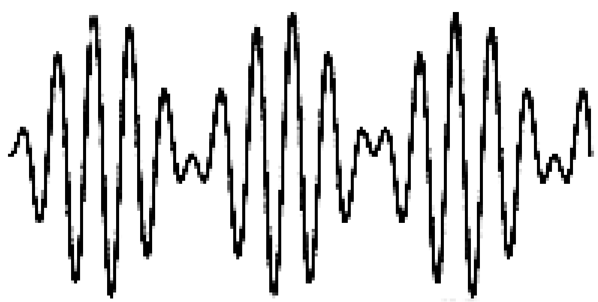
eg mixture of Gaussians

Prior language model

eg unigrams

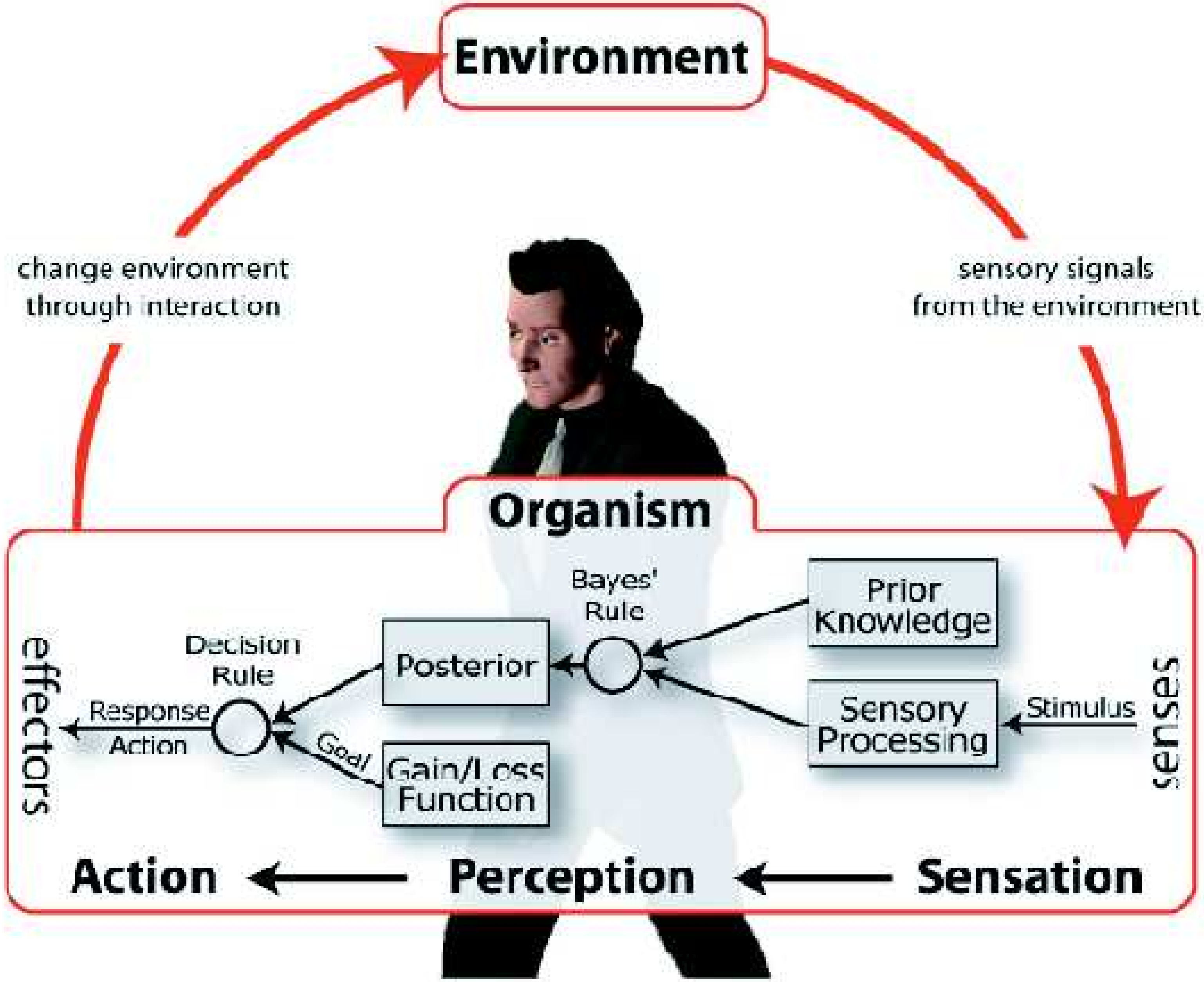
Hidden Markov Model (HMM)

“Recognize speech”



“Wreck a nice beach”

People as Bayesian reasoners



Bayes and decision theory

Utilitarian view: We need models to make the right decisions under uncertainty. Inference and decision making are intertwined.

Learned posterior

$$\begin{cases} P(x=\textbf{healthy}|data) = 0.9 \\ P(x=\textbf{cancer}|data) = 0.1 \end{cases}$$

Loss/Reward model *u(x,a)*

	$\mathbf{a} = \textit{no treatment}$	$\mathbf{a} = \textit{treatment}$
$\mathbf{x} = \textit{healthy}$	0	-30
$\mathbf{x} = \textit{cancer}$	-100	-20

We choose the action that maximizes the **expected utility**:

$$EU(a) = \sum_{x} u(x,a)\, P(x|data)$$

$EU(a=\textbf{treatment}) =$ u(treatment, healthy) 0.9 + u(treatment, cancer) 0.1
$= (-30)(0.9) + (-20)(0.1) = -29$

$EU(a=\textbf{no treatment}) =$ $(0)(0.9) + (-100)(0.1) = -10$

Don't treat

Next lecture

In the next class, we will introduce probabilistic graphical models.

These will enable us to attack problems with many variables.