

Lecture 11. Bayesian inference

How the brain infers the world?  
How we infer the happenings in brain?

- Bayes rule
- Bayesian inference
- Bayesian decoding of hippocampus
- Perception as Bayesian inference
- Illusion as Optimal Bayes

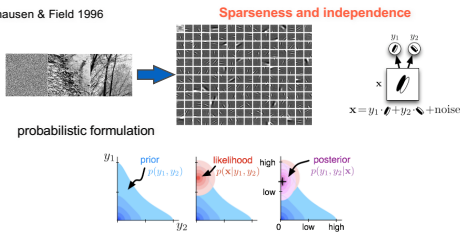
Bayes approach

- Alan Turing cracking the Enigma code.
- Bernard Koopman narrowing down the search for German U-boats – Bayesian update.
- John Craven finding the lost Hydrogen bomb in Palomares B-52 crash incident.

Susan McGrayne: The Theory That Would Not Die: How Bayes' Rule Cracked the Enigma Code, Hunted Down Russian Submarines, and Emerged Triumphant from Two Centuries of Controversy

Perception as an inference process based on evidence and priors

Olshausen & Field 1996



Bayes' Rule

$$P(H / E) = \frac{P(E | H)P_{prior}(H)}{\sum_{H'} P(E | H')P_{prior}(H')}$$

$H$  = a particular hypothesis  
 $E$  = evidence  
 $H'$  = all the hypotheses considered.

Bayes rule:

$$\text{Since } P(\omega_j, x) = P(\omega_j | x)P(x) = P(x | \omega_j)P(\omega_j)$$

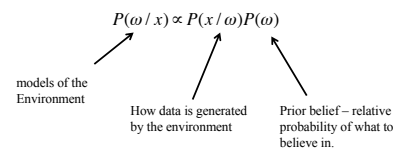
$$\text{posterior } P(\omega_j | x) = \frac{P(x | \omega_j)P(\omega_j)}{P(x)}$$

- where in case of two categories

$$P(x) = \sum_{j=1}^{c-2} P(x | \omega_j)P(\omega_j)$$

- Posterior (hypothesis given the data) = (
- Likelihood (data given the hypothesis) x Prior (hypothesis) / Prior(Observation)

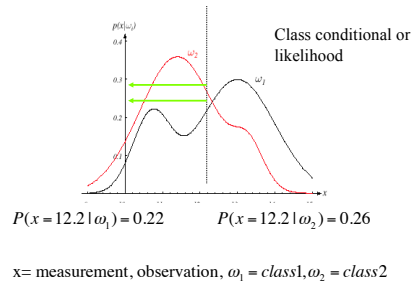
Bayes' Rule



If one event (catching Salmon) happens more often than the other (catching Chicken) ...

If  $P(\omega_1)=2/3$ , and  $P(\omega_2)=1/3$ , without making an observation, how would I guess?

Now, if the observation indicates that  $x=12.2$  inches, how would I decide? First find the likelihood  $p(x|w)$ .



### Bayes Decision (MAP) Rule

- Choose category  $i$  that has the maximum posteriori probability (MAP):

$$P(\omega_i | x)$$

- This produces the minimum probability of error:

$$p_e = 1 - P(\text{chosen category} | x)$$

Probability of making ALL the wrong choices.



Thomas Bayes 1702-1761

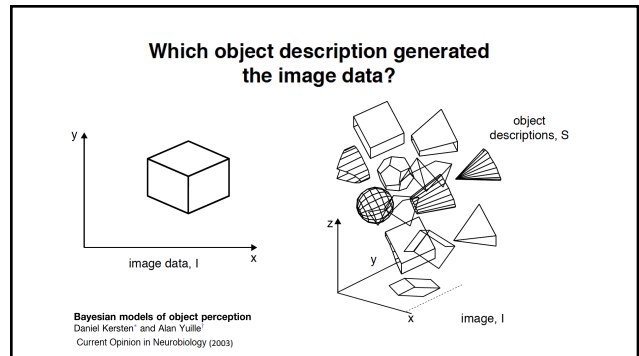
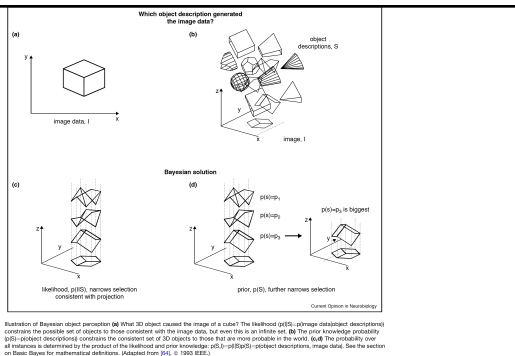
$$P(x) = \sum_{j=1}^{j=2} P(x | \omega_j) P(\omega_j)$$

$$= \frac{2}{3} \times 0.22 + \frac{1}{3} \times 0.26 = 0.233$$

$$P(\omega_1 | x = 12.2) = \frac{P(x = 12.2 | \omega_1) P(\omega_1)}{P(x = 12.2)} = \frac{0.1466}{0.233} = 0.629$$

$$P(\omega_2 | x = 12.2) = \frac{P(x = 12.2 | \omega_2) P(\omega_2)}{P(x = 12.2)} = \frac{0.0866}{0.233} = 0.371$$

Which one would you choose?  
What is the probability of error?



### Which object description generated the image data?

likelihood,  $p(I|S)$ , narrows selection consistent with projection

prior,  $p(S)$ , further narrows selection

Current Opinion in Neurobiology

### Motion Perception

Assuming small motion, the change in intensity is equal to the movement in the image gradient.

$$\frac{\partial I}{\partial x}u + \frac{\partial I}{\partial y}v = -\frac{\partial I}{\partial t} \quad \text{using } u = \frac{dx}{dt}, v = \frac{dy}{dt}$$

Optical flow constraint

$$D(r) = \sum_{x,y} \left( \frac{\partial I}{\partial x}u + \frac{\partial I}{\partial y}v + \frac{\partial I}{\partial t} \right)^2$$

### Velocity Space

- The optical-flow constraint: the flow velocity  $(u,v)$  has to lie along a straight line perpendicular to the brightness gradient.

$$\frac{\partial I}{\partial x}u + \frac{\partial I}{\partial y}v = -\frac{\partial I}{\partial t}$$

$$(I_x, I_y) \cdot (u,v) = -I_t$$

- We can determine the component in the direction of the brightness gradient, but not the flow components in the direction at right angles.

### Resolving ambiguity in local motion estimates

Show aperture movie

- How to integrate local cues?
  - intersection of constraints (IOC)
  - vector averaging (VA)
  - "blob" tracking

Show Yair Weiss Rhombus movie

### What are the priors in motion perception?

Given the same evidence, do we prefer an interpretation rest on slow motion or fast motion? Do things tend to move fast or move slowly?

Show Ellipse.

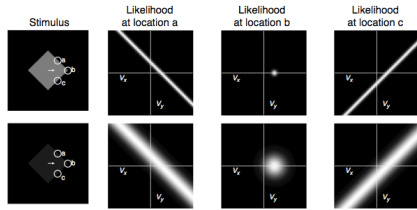
Priors favoring slow speed:

$$P(v) \propto \exp(-||v||^2/2\sigma_v^2)$$

The posterior probability of a velocity was computed by combining the likelihood and prior using Bayes' rule. Because we assumed that the noise is independent over spatial location, the total likelihood function is just a product of likelihoods:

$$P(v|I) \propto P(v) \prod_i P(I(x_i, y_i, t) | v), \quad (6)$$

How to encode uncertainty about the constraints?

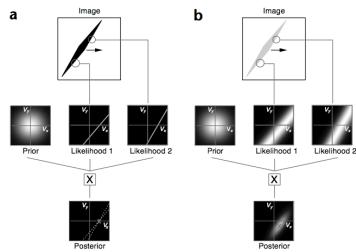


Posterior on velocity

$$P(v|I) \propto \exp \left( -\|v\|^2/2\sigma_p^2 - \frac{1}{2\sigma^2} \int_{x,y} \sum_i w_i(x,y) (I_x(x,y)v_x + I_y(x,y)v_y + I_i)^2 dx dy \right)$$

Here we assumed the entire image moves according to a single translational velocity, and so summed over all spatial positions. In this case,  $\sum_i w_i(x,y)$  is a constant, so the posterior probability is given by:

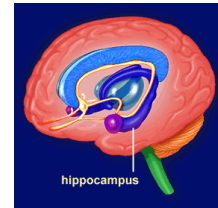
$$P(v|I) \propto \exp \left( -\|v\|^2/2\sigma_p^2 - \frac{1}{2\sigma^2} \int_{x,y} (I(x,y)v_x + I_y(x,y)v_y + I_i)^2 dx dy \right)$$



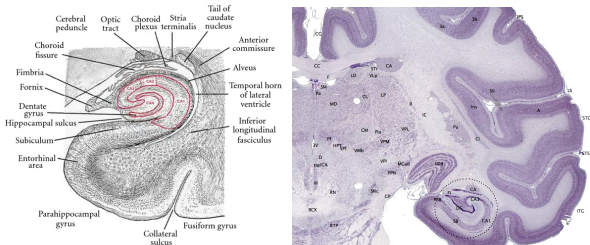
Yair Weiss<sup>1</sup>, Eero P. Simoncelli<sup>2</sup> and Edward H. Adelson<sup>3</sup> **Motion Illusions as optimal percepts**  
*nature neuroscience* • volume 5 no 6 • June 2002

Hippocampus

Limbic system structure important in forming new memories and connecting emotion and senses (smell, sound) to memories. Primitive brain structure located on top of the brain stem, buried under the cortex, involved in emotion and motivation, feelings of pleasures (eating and sex).

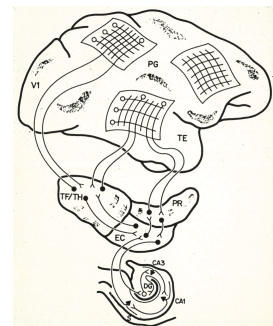


Hippocampus



Hippocampus

- Connectivity
  - receives input from much of cortex as well as subcortical structures
  - projects back to these same structures
  - recurrent network in HC, what is it good for?



Place cells in the CA1, CA3 and dentate gyrus in rodent hippocampus

The diagram illustrates the process of identifying place cells in the rodent hippocampus. It shows a cross-section of the hippocampus (CA1, CA3, and dentate gyrus), a rat in a circular arena, and a 3D reconstruction of the rat's path. Below, a sequence of heatmaps shows the firing rate of a single place cell, transitioning from a sparse distribution to a clear, localized peak in a specific area of the arena.

How to decode the location of the rat based on the ensemble activities?

This diagram shows a rat's path in a circular arena, with different colors representing different locations. A large arrow points from the path to a vertical bar representing an ensemble activity vector, where each color corresponds to the firing rate of a specific place cell at that location.

Zhang et al., 1998, Journal of Neurophysiology

Neural decoding:

The diagram illustrates the process of neural decoding. It shows two cells (cell 1 and cell 2) with their respective firing rate tuning curves. The tuning curves are plotted against location (T1T2T3T4 and F1F2). The probability distributions for each cell are also shown, representing the likelihood of the rat being at a specific location given the cell's firing rate. The final step shows the combined evidence from both cells, resulting in a more precise probability distribution for the rat's location.

cell 1  
cell 2

what location do these spikes represent?  
(where is the rat?)

cell 1 tuning curve  
cell 2 tuning curve

probability  
probability

Combine evidence from two cells

Naïve Bayes

- Bayes rule
 
$$P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)}$$

$$X = (X_1, \dots, X_N)$$
- Naïve Bayes assumes that  $X_i$  are conditionally independent, given  $Y$ 

$$P(X|Y) = P(X_1, \dots, X_N|Y) = \prod_{i=1}^N P(X_i|Y)$$

Algorithm

- Step 1: Create each cell's tuning curve
- Occupancy:
 
$$P(x) = \frac{N(x)}{\sum_x N(x)}$$
- Tuning curve:
 
$$f(x) = \frac{S(x)}{N(x)\Delta t}$$

$N(x)$  = number of discrete time steps the rat spend at each location.  
 $\Delta t$  = the duration (ms) of each discrete time step.  
 $f(x)$  = average spike counts per unit time at each location.  
 $S(x)$  = spike counts at location  $x$ .  
 $n$  = number of spikes within a time window.

Tuning curve re-expressed

$$f(x) = \frac{S(x)}{N(x)\Delta t}$$

$$f_{mean} = \frac{\sum_x S(x)}{\sum_x N(x)\Delta t}$$

$$P(x|''spike'') = \frac{S(x)}{\sum_x S(x)}$$

$$P(x) = \frac{N(x)}{\sum_x N(x)}$$

$$f(x) = f_{mean} \frac{P(x|''spike'')}{P(x)}$$

### Bayes

$$f(x) = f_{mean} \frac{P(x | \text{spike}^n)}{P(x)}$$

$$P(x|\mathbf{n})P(\mathbf{n}) = P(\mathbf{n}|x)P(x)$$

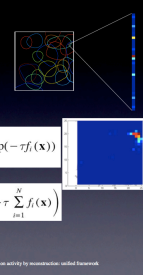
$$P(\mathbf{n}|x) = \prod_{i=1}^N P(n_i|x) = \prod_{i=1}^N \frac{(\tau f_i(x))^{n_i}}{n_i!} \exp(-\tau f_i(x))$$

$$P(x|\mathbf{n}) = C(\tau, \mathbf{n}) P(x) \left( \prod_{i=1}^N f_i(x)^{n_i} \right) \exp\left(-\tau \sum_{i=1}^N f_i(x)\right)$$

$$x_{Bayes} = \underset{x}{\operatorname{argmax}} P(x|\mathbf{n})$$

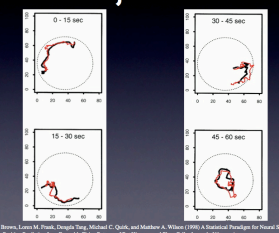
$f_{mean}$  = mean firing rate of the neurons across all locations.

$n$  = number of spikes within a time window.



Zhang & Glimcher, J. Neurophysiol. 93: 533-541 (2005) encoding normal population activity by non-linear, linked networks with sparse & hierarchical structure. (Neuron 54: 931-943)

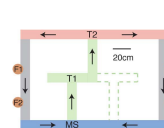
### Decoding Location/ Trajectories



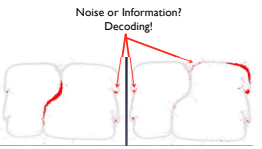
Ernst & Bressan, Nature 415: 429-433 (2002) A Statistical Perspective for Visual Spike Train Decoding Applied to Position Prediction from Unsteady Firing Patterns of the Dorsocentral Parietal Cells. (Journal of Neurophysiology)

### Is there replay during awake?

- During sleep and inattentive pauses in the environment, animals replay sequences of neural activity representing past experiences. Thought to reflect the process of consolidation.



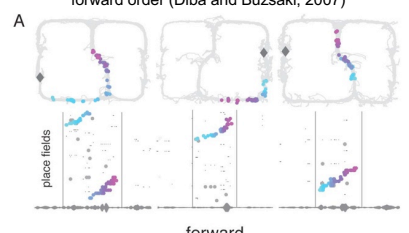
Noise or Information?  
Decoding!



### Is there replay during awake?

What is about to be experienced in the forward order (Diba and Buzsaki, 2007)

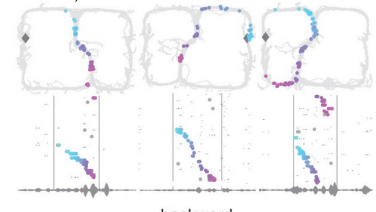
A



forward

### Is there replay during awake?

What was just experienced in the reverse order (Foster and Wilson, 2006)

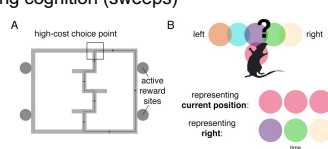


backward

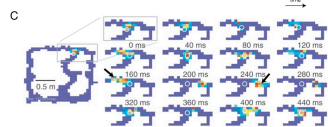
### Decoding cognition (sweeps)

A high-cost choice point

B



C



David Redish