The `Bayesian Brain'

Readings:Knill & Pouget, *TINS*, 2004

Uncertainty everywhere

- Humans and other animals operate in a world of sensory uncertainty:

- e.g. mapping of 3D objects to 2D image
- intrinsic limitations of the sensory systems (e.g. number and quality of receptors in the retina)
- neural noise

--> multiple interpretations about the world are possible;

- The brain must deal with this uncertainty to generate perceptual representations and guide actions.

- Perception as unconscious, probabilistic inference

Example: Multi-sensory integration

- How to integrate information coming from different senses?

- example: ventriloquism.

- sound has uncertain origin.
- visual image is that the puppet is talking.

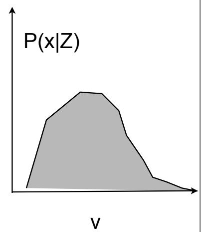
- at which stage are the informations integrated to provide estimate of spatial origin of the sound ? how?

Bayesian coding hypothesis (1)

- Hypothesis: information provided by sensory systems has the form of a conditional probability density function

- e.g. the position of an object is not a single number, x, but P(x|Z), where Z is the available data

- = stores likelihoods = 'generative models', or 'forward model' of the world, P(Z|x), and prior knowledge / state of the world, P(x).
- Given new data Z, the brain computes P(x|Z)



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

Bayes theorem

Bayesian coding hypothesis (2)

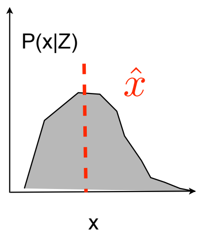
- Benefits:

\- integrate information efficiently over space & time

\- integrate information efficiently from different sensory cues and modalities

\- propagate information without committing too early to particular interpretations.

- Commit as late as possible, then collapsing the probability distribution into a single number = decision, or action taken.
- How to do that depends on cost function :



$$\hat{x} = \text{argmax}_x P(x|Z)$$ max of the posterior

Are Humans Bayes - optimal?

- Humans not optimal / achieving the level of performance afforded by the uncertainty in the physical stimulus (e.g. movies)

- The question is:

Do the neural computations take into account the uncertainty at each stage of processing?

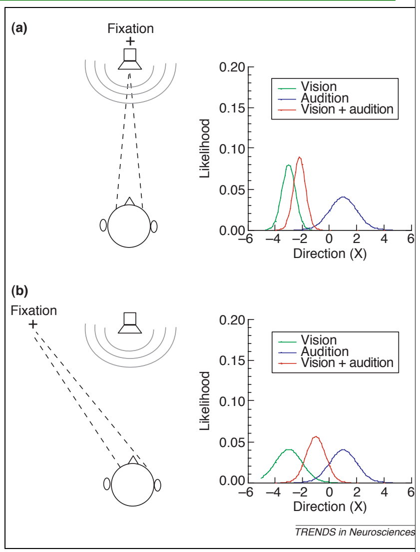
- Bayesian hypothesis makes a lot of testable predictions on how different sources of uncertainty should be integrated. Valid?

Cue Integration (1) : qualitative predictions

- e.g. integration between visual and auditive information

- prediction 1 (position): if visual cue is more reliable, then final estimate is shifted towards visual cue.

- prediction 2 (variance or discrimination threshold): Final discrimination threshold lower than that for each modality ; varies if reliability of one modality varies.



Cue Integration (2):Theory

- Theory tells us how posterior depends on individual likelihoods:

$$\hat{x} = \text{argmax}_x P(x|d_1, d_2)$$

$$P(x|d_1, d_2) = \frac{P(d_1, d_2|x)P(x)}{P(d_1, d_2)} \propto P(d_1|x)P(d_2|x)P(x)$$

- Assuming that the likelihood are gaussian, i.e.

$$P(d_1|x) \propto \exp(-\frac{(d_1 - x)^2}{2\sigma_1^2})$$

- We can determine mean and width of posterior (gaussian):

$$P(d_1|x)P(d_2|x) \propto \exp(-\frac{(d_1 - x)^2}{2\sigma_1^2} - \frac{(d_2 - x)^2}{2\sigma_2^2}) \propto \exp\frac{\left[x - \frac{\sigma_2^2 d_1 + \sigma_1^2 d_2}{\sigma_1^2 + \sigma_2^2}\right]^2}{2\sigma_1^2\sigma_2^2/(\sigma_1^2+\sigma_2^2)}$$

Cue Integration (3):Theory

- If we know mean estimate and variance for each modality in isolation, we can deduce mean of bimodal estimate:

$$\mu = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} d_1 + \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} d_2$$

pushed towards more reliable cue

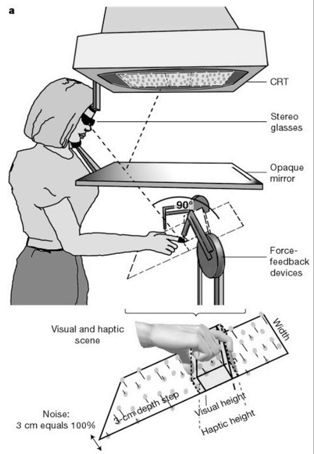
- and discrimination threshold

$$T_{1,2}^2 \propto \sigma_{1,2}^2 = \sigma_1^2 \sigma_2^2 / (\sigma_1^2 + \sigma_2^2)$$

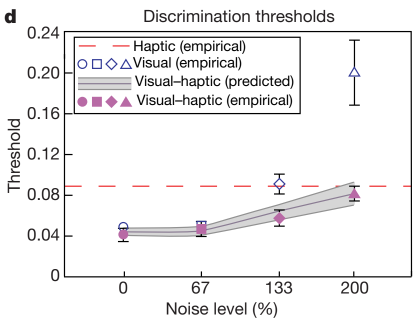
smaller than 1 or 2 alone

Cue Integration (4): Ernst and Banks, *Nature*, 2002

- visual + haptic cues
- vary noise level / visual cue
- compute discrimination threshold for each cue alone, or when both are present.



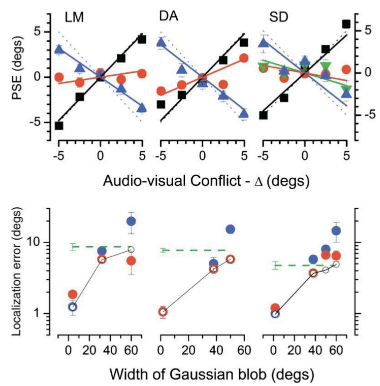
Cue Integration (5): Ernst and Banks, *Nature*, 2002



- height jugment follows optimal integration of visual and haptic cues.
- 'visual capture' for low visual noise
- 'haptic capture' for high visual noise
- instanteous 'switch'
- numerous studies replicate this result in a variety of paradigms.

Cue Integration (6):Ventriloquist effect

- Alais & Burr, *Curr Biol*, 2004
- visual blob of various size + auditive 'click', possibly in conflict.
- measure both estimate of position (mean), and discrimination threshold
- near optimal integration
- visual capture for small blobs
- auditive capture for large blobs

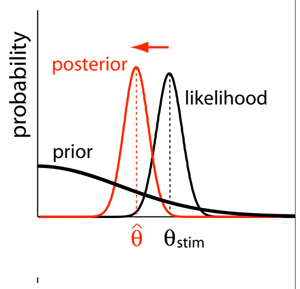


Other predictions of Bayesian theory

• Prior knowledge about the world can be used to interpret data in situation of uncertainty.

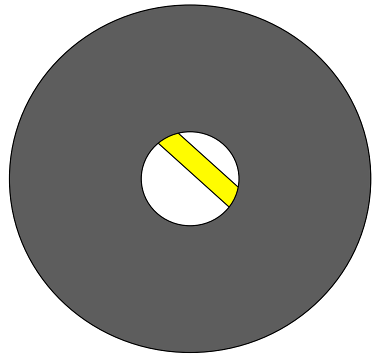
• Prediction: the more uncertain the data, the more the prior should influence the interpretation.

• The priors should reflect the statistics of the sensory world.

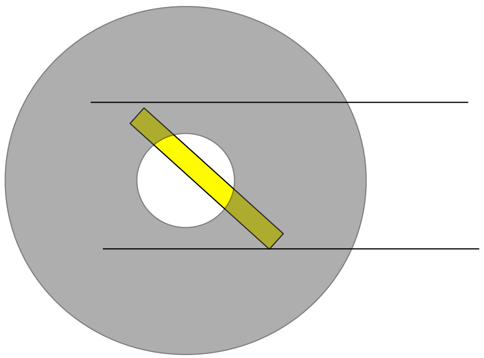


Interpreting motion (1): the aperture problem

• what is the direction of the line?



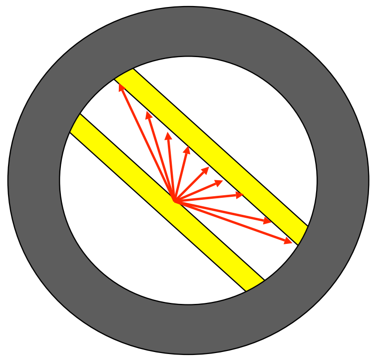
Interpreting motion (1): the aperture problem

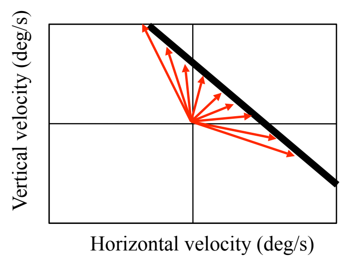


Interpreting motion (2): the aperture problem

• Motion shown in an aperture is fundamentally ambiguous; it can be interpreted in an infinite number of ways

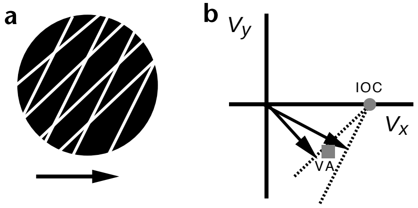
• which one is chosen? why?





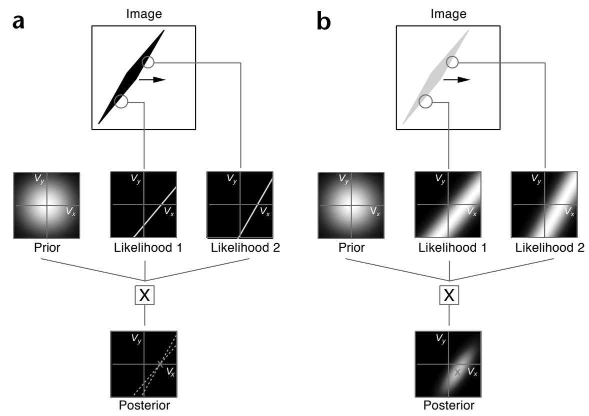
Interpreting motion (3): the aperture problem

- More complex stimuli can be constructed, by adding more segments of ambiguous motions, e.g. plaid, or rhombus.
- How is the system going to integrate the different possible interpretations?
- classical models: intersection of contraints (IOC), Vector Averaging (VA), feature tracking.
- donnot capture the complexity of available data.



Interpreting motion : A Prior on Low Speeds (1)

- Hypothesis: humans tend to favor slower motions
- Use a (gaussian) prior on low speeds (centered at 0).



Weiss, Adelson & Simoncelli, *Nat Neuro*, 2002

interpreting motion : A Prior on Low Speeds (2)

- provides a very simple model which explains a large variety of psychophysical effects / perception of plaids, rhombus and plaids, barber pole and effects of contrast [Weiss et al, 2004]
- Thomson effect: humans tend to underestimate speed at low contrast (why drivers tend to speed up in the fog)
- Stocker & Simoncelli (2005) measure the shape of the prior.
- illusions as 'optimal percepts'.

- .... elegant psychophysics looking for neural basis.

Neural implementation ?

- How do populations of neurons represent uncertainty ?
- Does neural activity represent probabilities? (log probabilities?)
- Can we distinguish stages where the likelihoods, priors, posterior could be ‘measured’ experimentally ?
- Can networks of neurons implement optimal inference?
- How can we discover the priors used by the brain?
- How can a prior be implemented? (baseline - spontaneous activity, number of neurons, gain, connectivity?).
- Recently, active topic of theoretical research (A. Pouget, S. Deneve, P. Dayan, R. Rao).
- promising direction for PhD project :-)

Thanks !

Bayes’ Theorem

- Bayes’ theorem is a result in probability theory that relates conditional probabilities P(AlB) and P(BlA)
- Given the likelihood and the prior, we can compute the posterior.

$$P(h_1|e) = \frac{P(e|h_1)P(h_1)}{P(e)}$$

$$\text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{normalizing constant}}$$

Reverend Thomas Bayes, 1702- 1761