02458 Cognitive Modeling

Bayesian Models of Visual Perception

Kersten and Yuille

Conditional probability

General Formula for $P(A|B)$

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

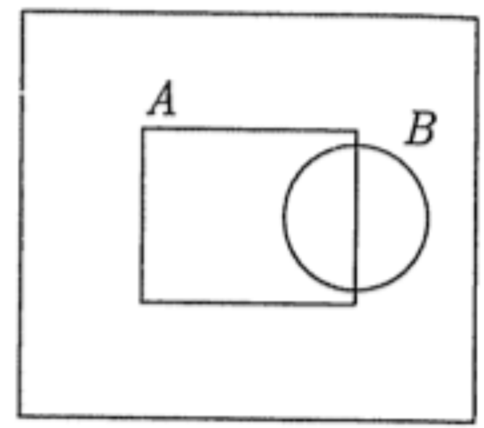
Conditional probability

Relative areas.

Suppose a point is picked uniformly at random from the big rectangle in the diagram. Imagine that information about the position of this point is revealed to you in two stages, by the answers to the following questions:

Question 1. Is the point inside the circle B?

Question 2. Is the point inside the rectangle A?



If the answer to Question 1 is yes, what is the probability that the answer to Question 2 will be yes?

The problem is to find the probability that the point is in the rectangle A given that it is in the circle B. By inspection of the diagram, approximately half the area inside B is inside A. So the required probability is

$$P(A|B) = \frac{P(AB)}{P(B)} = \frac{\text{Area}(AB)}{\text{Area}(B)} \approx 1/2$$

Conditional probability

General Formula for $P(A|B)$

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

Multiplication Rule

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

$$P(BA) = P(B \mid \mathrm{A})P(A)$$

Bayes' rule

$$P(A \mid \mathrm{B})P(B) = P(AB) = P(B \mid \mathrm{A})P(A)$$

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

What’s the fuzz?!

$$P\left(Hypothesis \mid \mathrm{data}\right) \propto P\left(Data \mid Hypothesis\right) P\left(Hypothesis\right)$$

- The problem of experimental science
  - EEG source localization
  - Geophysicist searching for oil
  - “Inverting probabilities” using
    - A generative model
    - Prior assumptions

What's the fuzz?!

- The problem of any observer
- The problem of perception

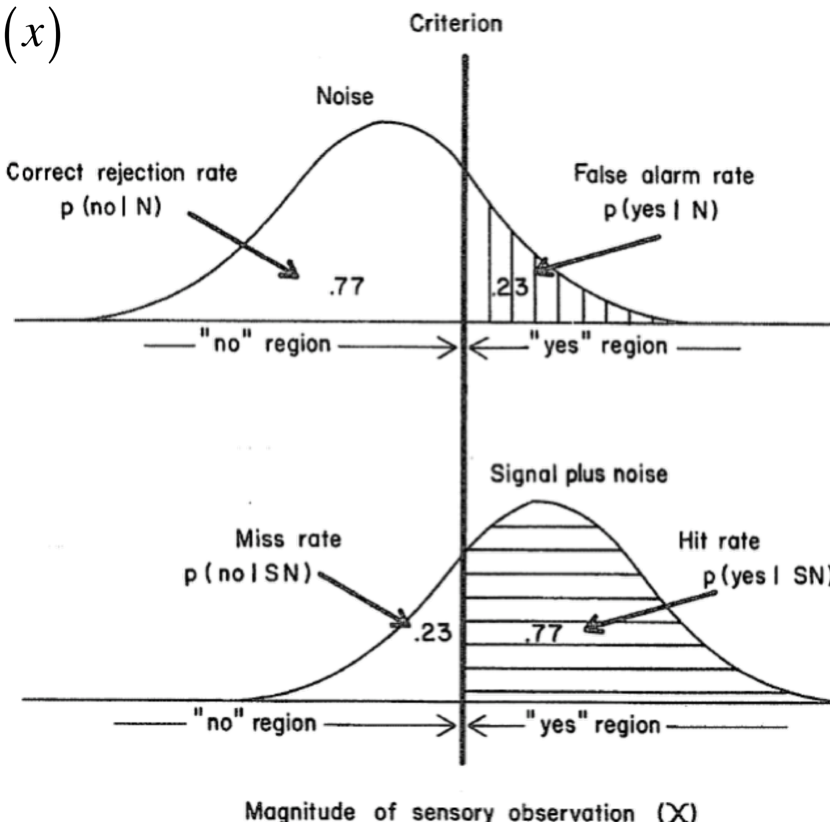
$$P(\text{the world} \mid \text{sensory stimulation}) \propto$$
$$P(\text{sensory stimulation} \mid \text{the world})P(\text{the world})$$

Bayes vs. Signal Detection Theory

Posterior probability densities:

$$P(noise \mid x) = \frac{P(x \mid noise)P(noise)}{P(x)} = \frac{\varphi(x)P(noise)}{P(x)}$$

$$P(signal \mid x) = \frac{P(x \mid signal)P(signal)}{P(x)} = \frac{\varphi(x-d')P(signal)}{P(x)}$$



Bayes vs. Signal Detection Theory

Max Aposteriori Rule: Yes if

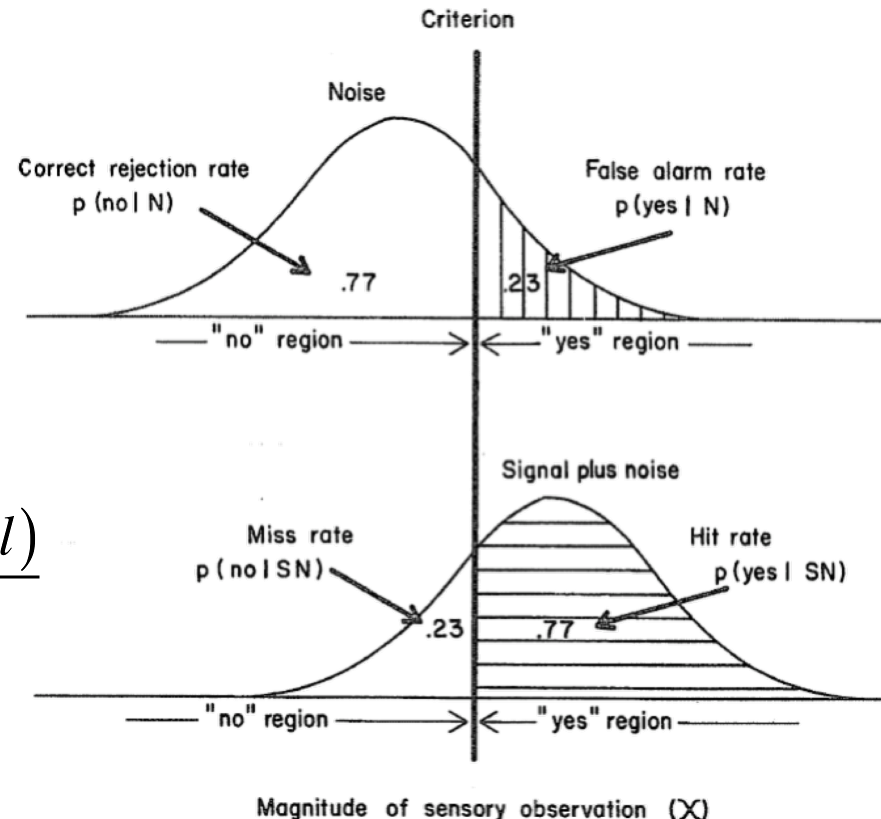
$$P(signal \mid x) > P(noise \mid x)$$

$$\frac{P(signal \mid x)}{P(noise \mid x)} > 1$$

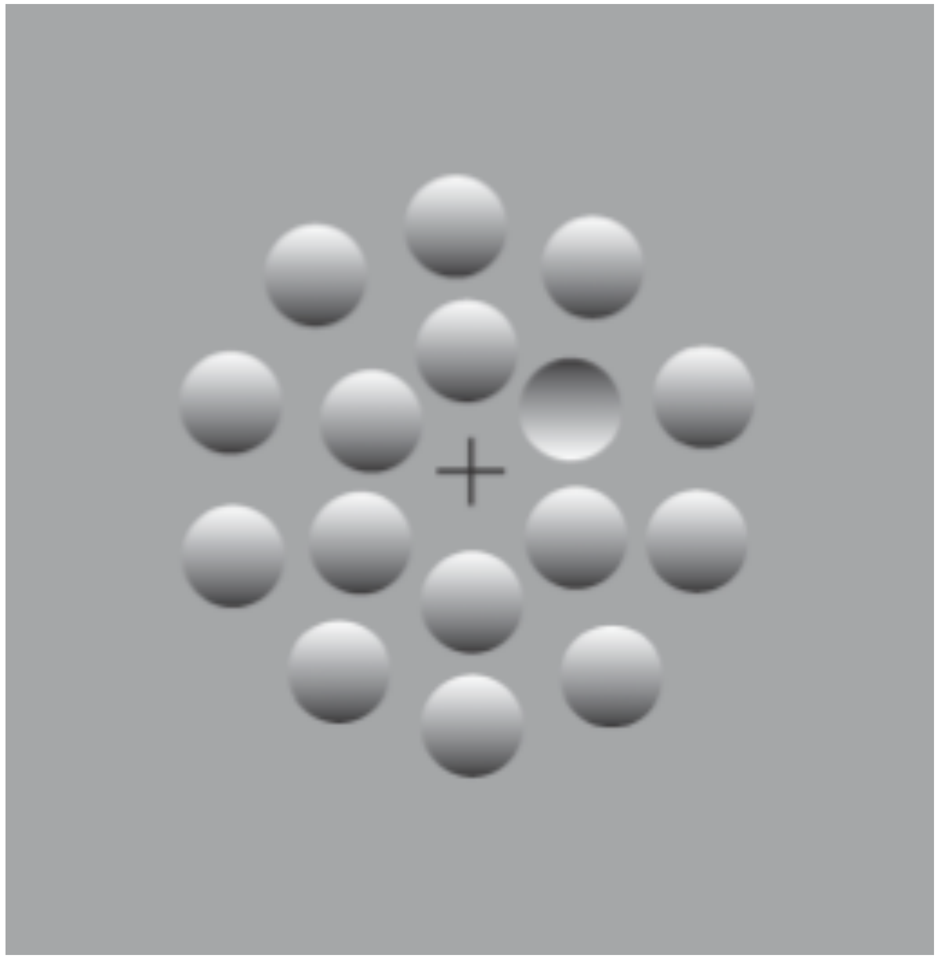
$$\frac{P(x \mid signal)\,P(signal)}{P(x \mid noise)\,P(noise)} > 1$$

$$\frac{P(x \mid signal)}{P(x \mid noise)} > \frac{P(noise)}{P(signal)} = \frac{1 - P(signal)}{P(signal)}$$

Defines the criterion

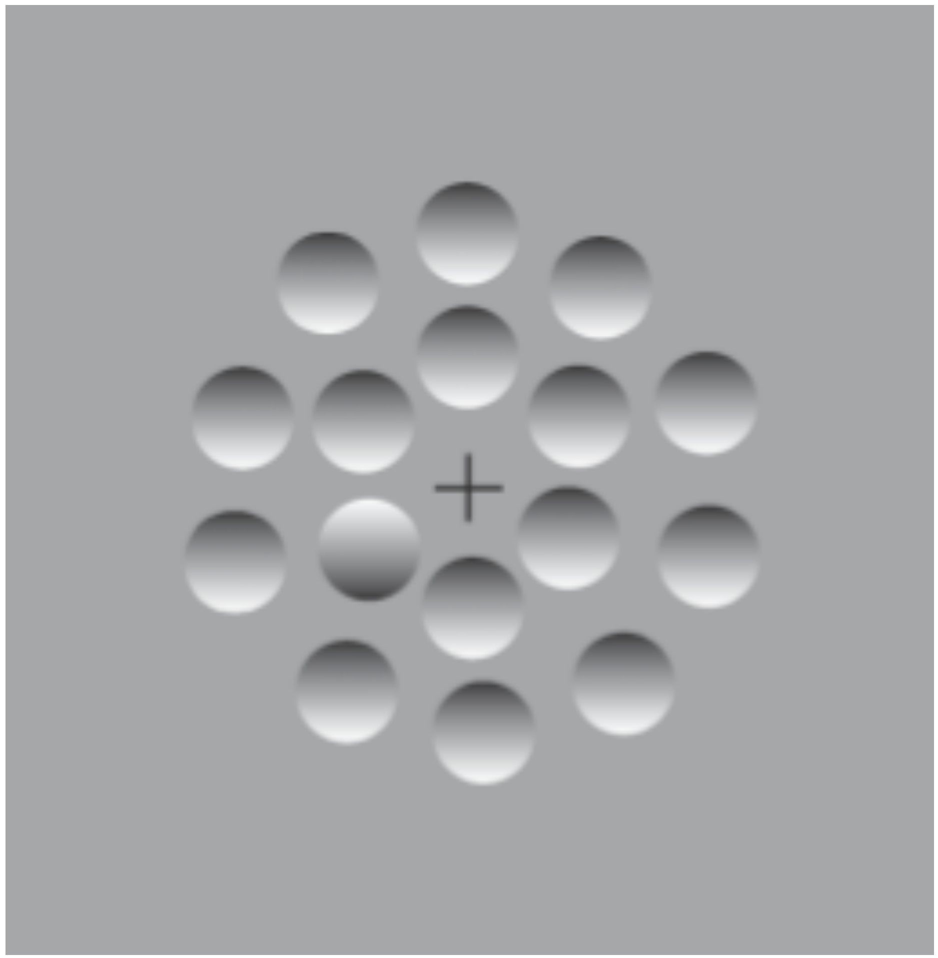


Protrusions or crevasses?

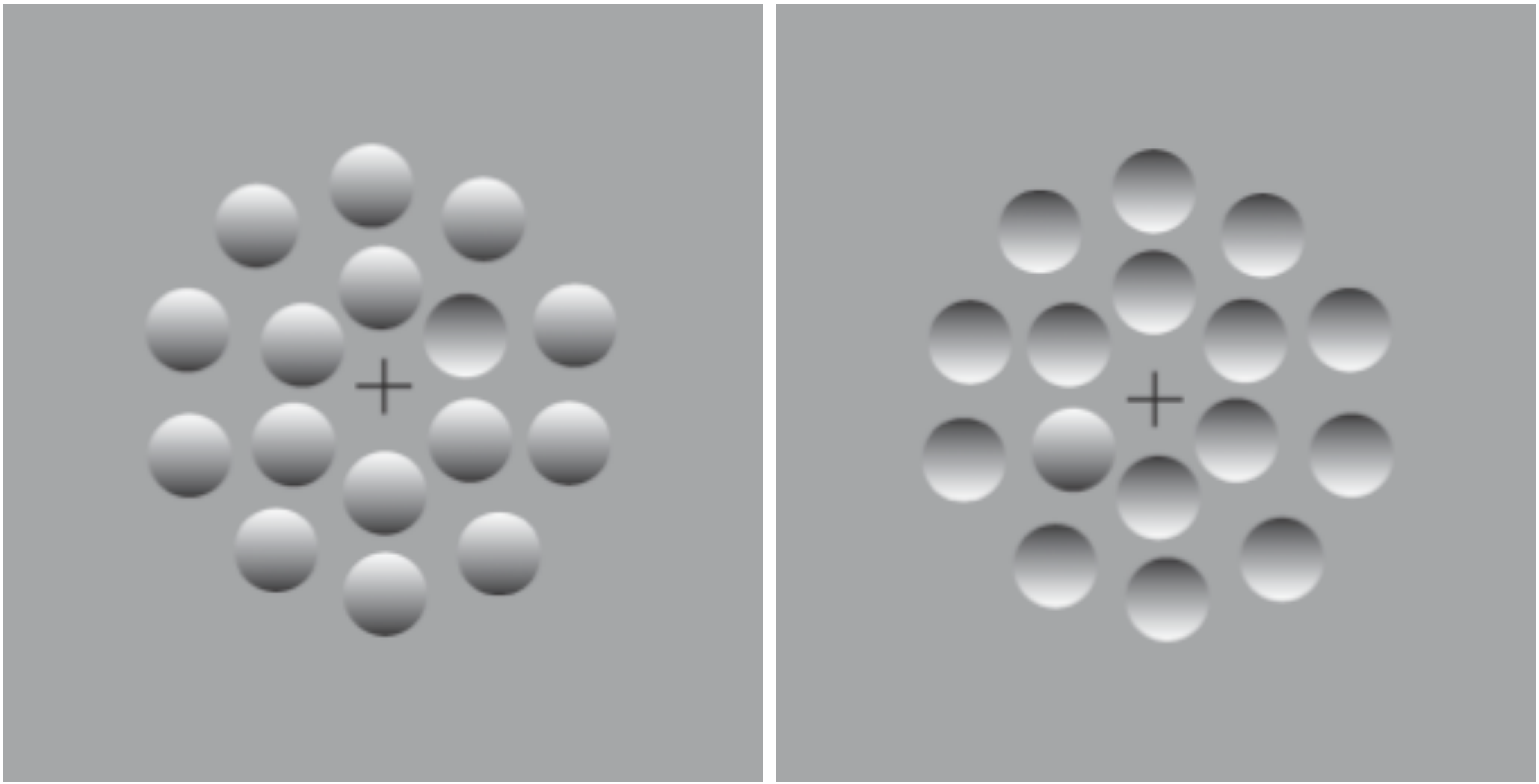


Champions and Adams, JOV, 2007

Protrusions or crevasses?

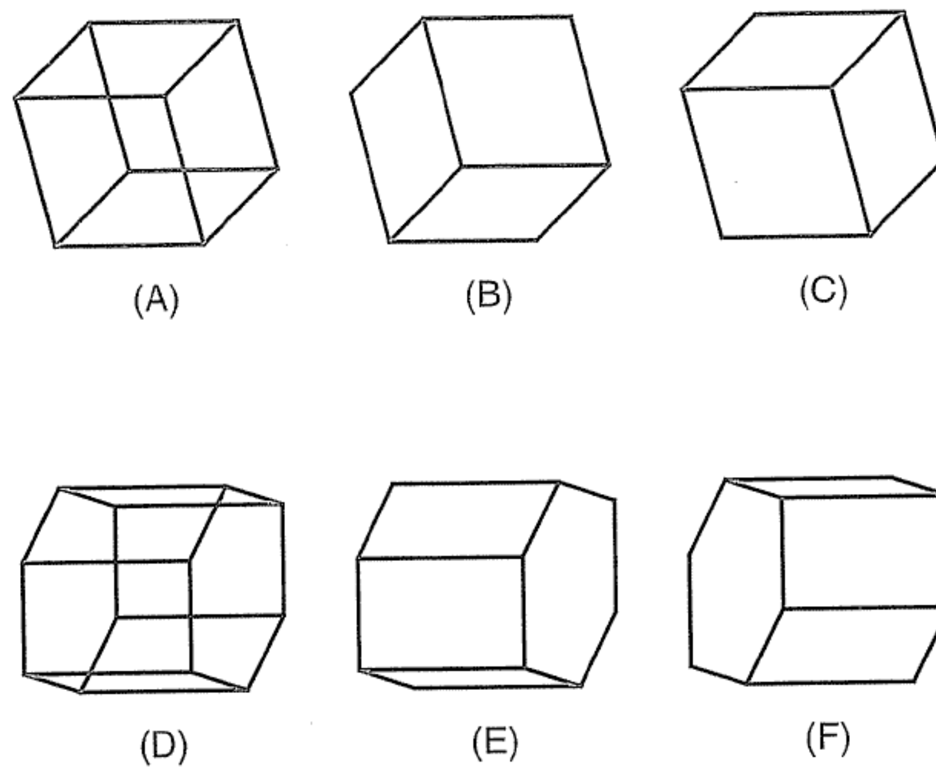


Protrusions or crevasses?



Bayes vs. Necker

- Vision is an ill-posed problem
- To solve it, we need additional info



Bayes vs. Necker

- Perception seems to go “beyond the information given”. This is seen when we perceive the Necker Cube as a cube and not some other 3-dimensional shape.
- How much information (in degrees of freedom) is in the Necker Cube?
- How much information is in a real cube?
- How much information perception adding? (Hint: Not just a simple subtraction)
- What general assumptions could perception be using in adding this information?

The (basic) Bayesian approach

- We get a 2D image, I
  - We want to infer a 3D scene, S, so

$$P(S \mid I) = \frac{P(I \mid S)P(S)}{P(I)}$$

  - where *P(I|S)* is the likelihood and *P(S)* is the prior

The (basic) Bayesian approach

- Let I be the image vector (dimension: 8×2 =16)
- Let S be the scene vector (dimension: 8×3 =24)
- The scene causes an image on the retina
  - Gaussian noise is added in the process
  - Gaussian prior

$$P\left(\hat{\mathrm{S}} \mid \mathrm{I}\right) \propto P(\mathrm{I} \mid \mathrm{S}) P(\mathrm{S})$$

$$P\left(\hat{\mathrm{S}} \mid \mathrm{I}\right) \propto \exp\left(-\frac{\left(\mathrm{I} - \text{projection of } \hat{\mathrm{S}} \text{ onto the image plane}\right)^2}{\sigma_{noise}^2}\right) \exp\left(-\frac{\left(\mathrm{S} - \hat{\mathrm{S}}\right)^2}{\sigma_{prior}^2}\right)$$

The (basic) Bayesian approach

- Decisions, decisions
  - We don't need the posterior $P(\mathrm{S} \mid \mathrm{I})$
  - We need a decision $\arg\max_S \left[P(\mathrm{S} \mid \mathrm{I})\right]$
- Arg max of a Gaussian?

$$\arg\max\left[\exp\left(-\frac{(\mathrm{x}-\mu)^2}{\sigma^2}\right)\right] = \arg\min\left[\frac{1}{\sigma^2}(\mathrm{x}-\mu)^2\right]$$

The (basic) Bayesian approach

- Arg max of the posterior?

$$\arg\min\left[\frac{1}{\sigma_{noise}^2}\left(I_{perceived}-\mathrm{I}_{projected}\right)^2+\frac{1}{\sigma_{prior}^2}\left(\mathrm{S}-\hat{\mathrm{S}}\right)^2\right]$$

- A weighted sum of squared errors
  - The likelihood error and the prior error(s)
- The product of two Gaussians is a Gaussian
  - The mean is the weighted sum
  - The *reliability* (inverse variance) sums

Decomposition of images and scenes

- Decomposing the
  - Image into features, I_1, I_2, …
    - Overall luminance
    - Wavelet decomposition
    - Binocular disparity
  - Scene into components, , S_1, S_2, …
    - Shape
    - Viewpoint
    - Illumination angle
  - Hard!

$$P\left(S_1, S_2 \mid I_1, I_2\right) = \frac{P\left(I_1, I_2 \mid S_1, S_2\right) P\left(S_1, S_2\right)}{P\left(I_1, I_2\right)}$$

Decomposition of images and scenes

- Arg max of the posterior?

$$\arg\min\left[\frac{1}{\sigma_{noise}^{2}}\left(I_{perceived}-\mathrm{I}_{projected}\right)^{2}+\frac{1}{\sigma_{prior}^{2}}\left(\mathrm{S}_{1}-\hat{\mathrm{S}}_{1}\right)^{2}+\frac{1}{\sigma_{prior}^{2}}\left(\mathrm{S}_{2}-\hat{\mathrm{S}}_{2}\right)^{2}+\ldots?\right]$$

- A weighted sum of squared errors
  - The likelihood error and the prior error(s)
    - Angles, location of light source, volume, etc.
  - Assuming that the scene components are independent

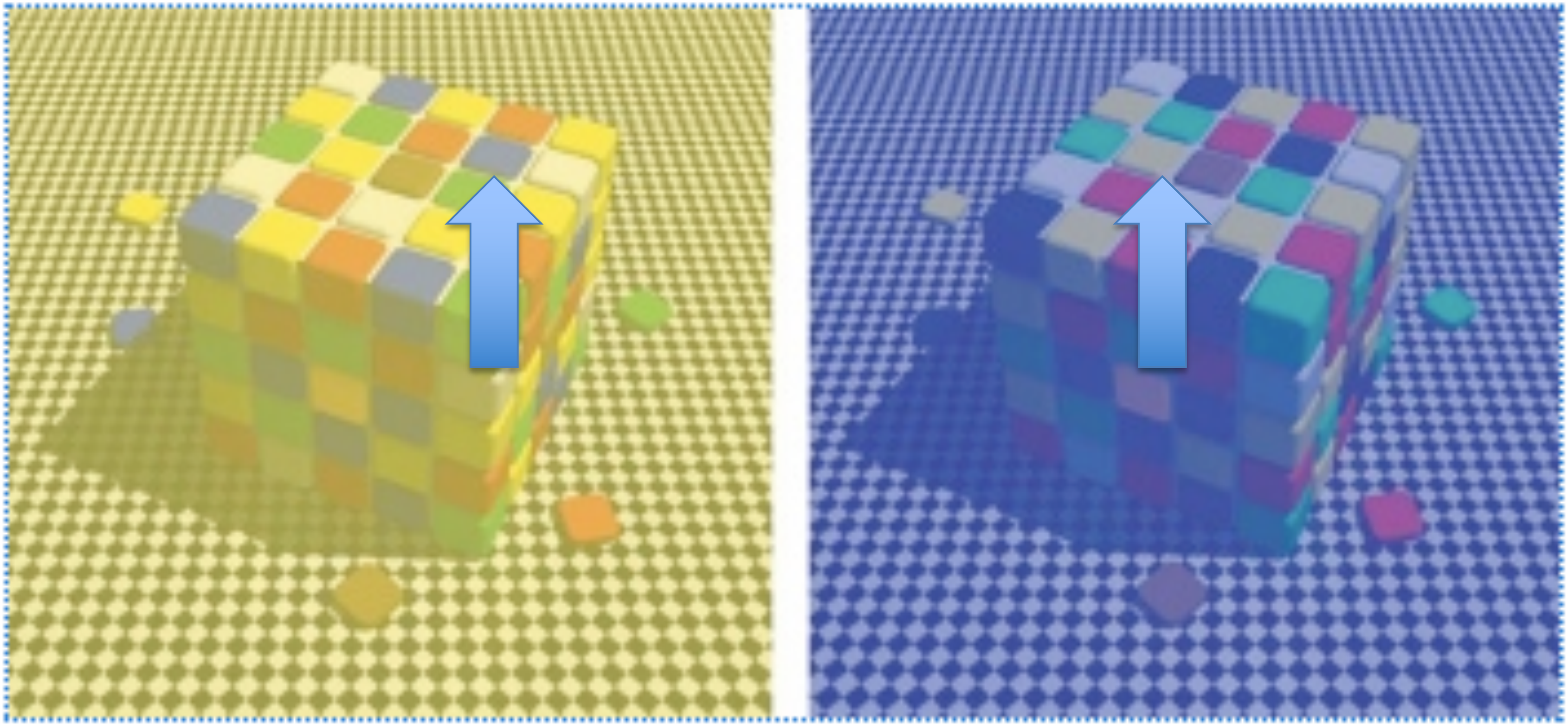
Discounting

- The task defines what is signal and what is noise
- Example: Find the bicycle
  - S_1: presence of bicycle
  - S_2: viewpoint
- Object constancy
  - Perceptual invariance

Discounting

Discounting

Discounting



Discounting

- We have a nuisance parameter
  - E.g. spectrum of the light source
- "Integrating out"
  - Or "summing out"
  - Is simply the weighted average of probabilities
  - Frees us from "unwanted" or "nuisance" parameters

$$P(A_1) = \int_{A_2} P(A_1, A_2) = \int_{A_2} \frac{P(A_1, A_2)}{P(A_2)} P(A_2) = \int_{A_2} P(A_1 \mid A_2) P(A_2)$$

Discounting

- What we got: $$P(S_1, S_2 \mid I) = \frac{P(I \mid S_1, S_2) P(S_1, S_2)}{P(I)}$$

- What we want:

$$P(S_1 \mid I) = \int_{S_2} P(S_1, S_2 \mid I) = \int_{S_2} \frac{P(I \mid S_1, S_2) P(S_1, S_2)}{P(I)}$$

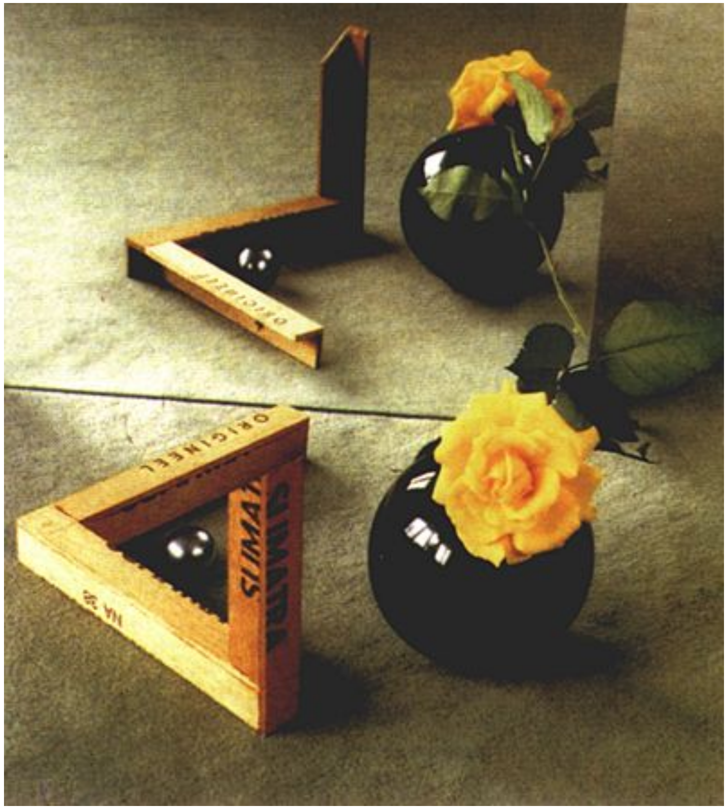
Discounting

- Assuming flat priors
  - And rearranging

$$P(S_1 \mid I) = \int_{S_2} P(S_1, S_2 \mid I) \propto$$

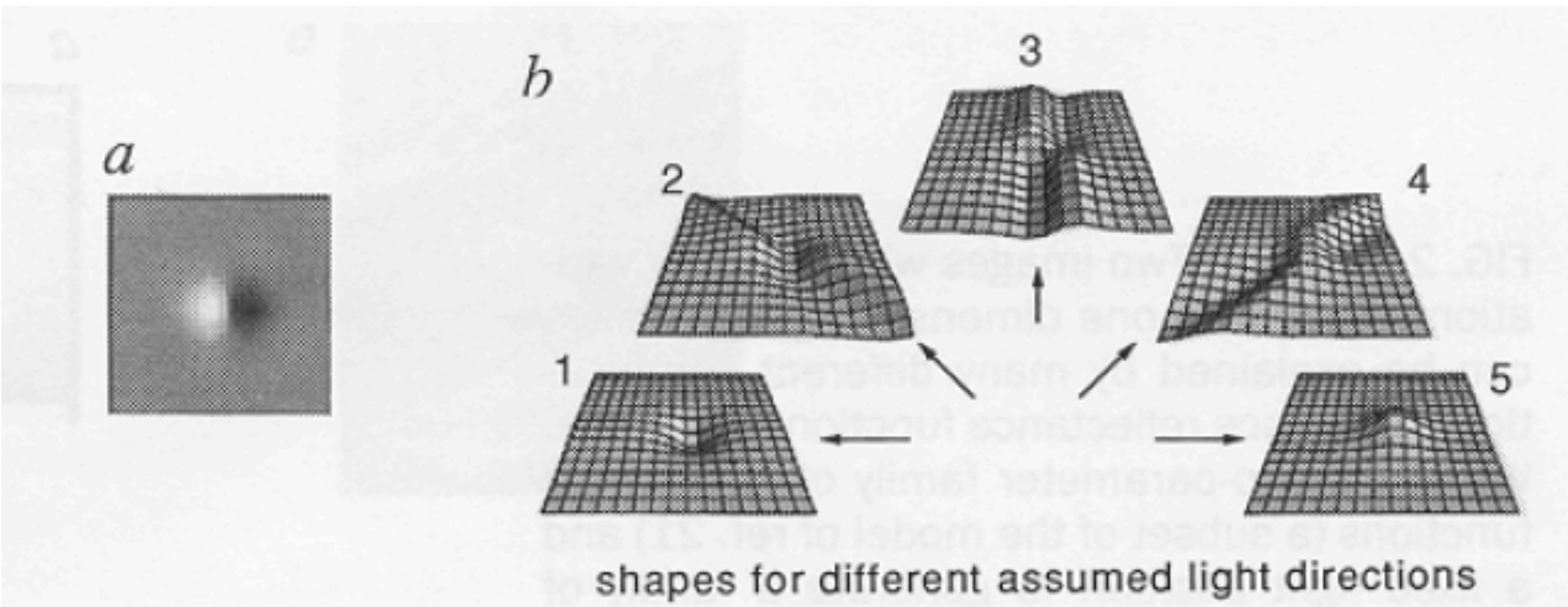
$$\int_{S_2} P(I \mid S_1, S_2) =$$

- Example: Necker cube
  - S_1 is the structure, S_2 is the viewpoint
  - Discounting = generic viewpoint assumption



Discounting

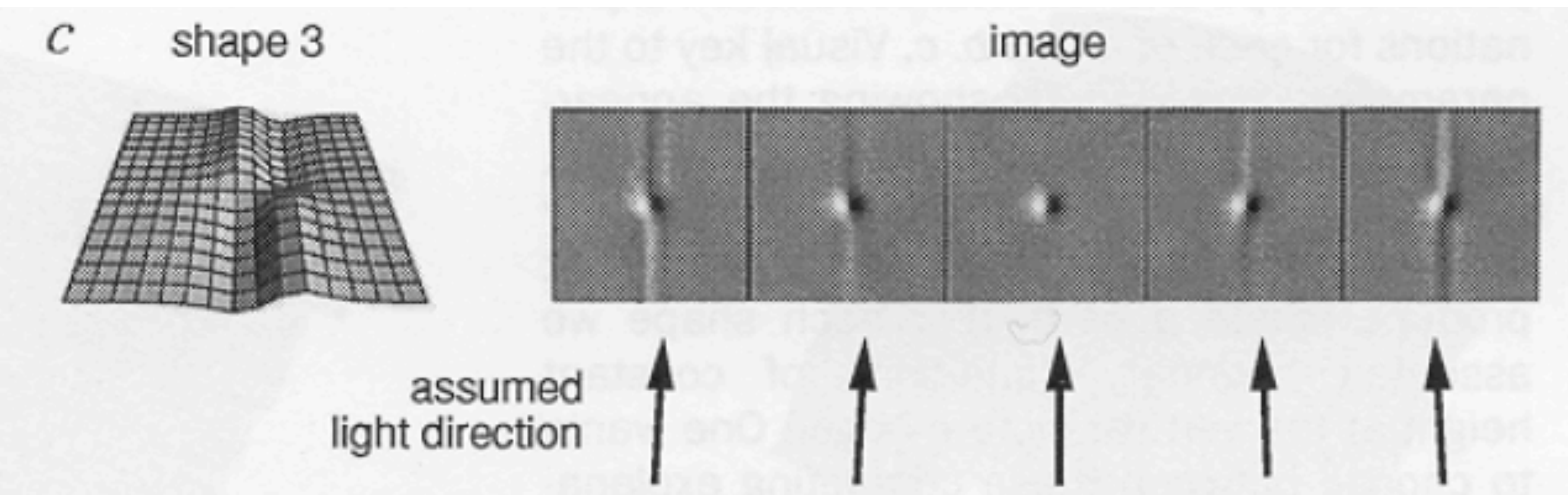
- Example: Light from above prior
  - Weird shapes possible



Freeman, 1994

Discounting

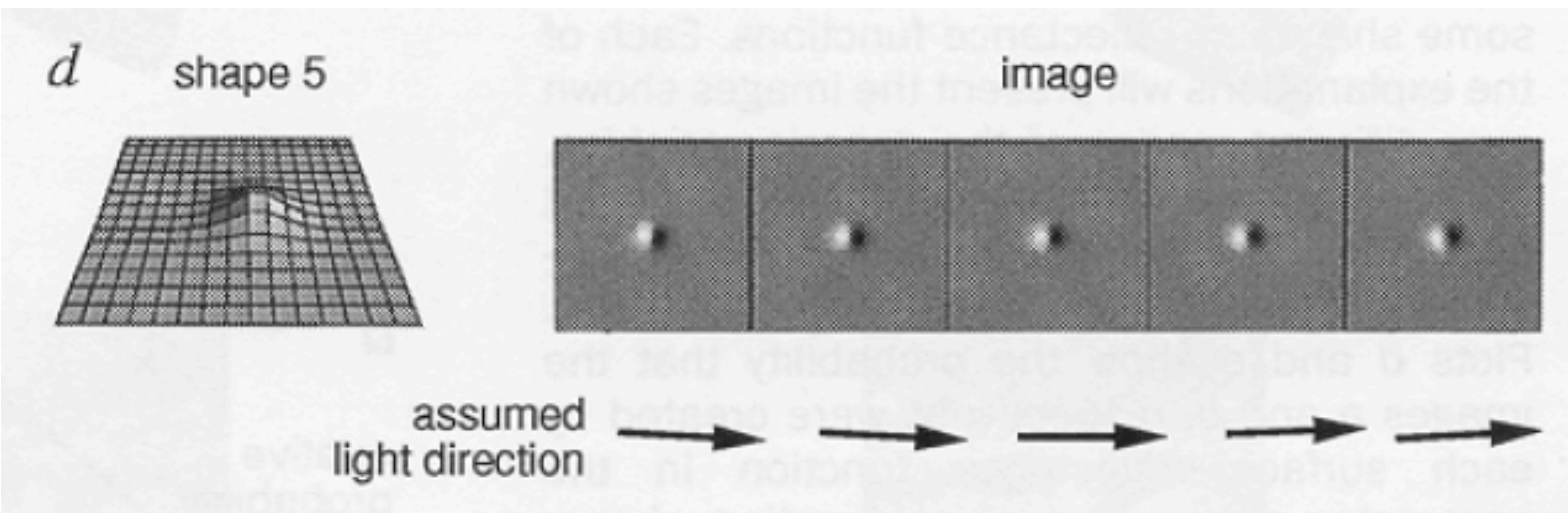
- Weird shapes very sensitive to illuminant angle



Freeman, 1994

Discounting

- Seen shapes not so sensitive to viewpoint

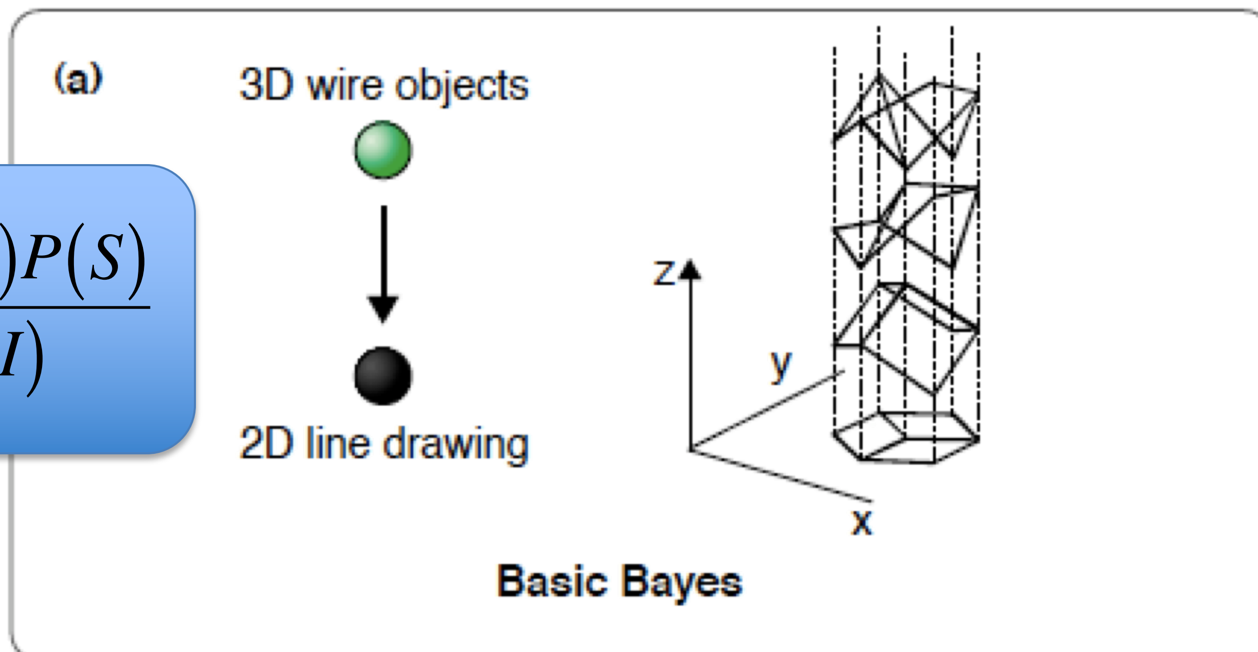


Freeman, 1994

Graphical models

- A simple representation of “causality”
- Directed Acyclic Graphs
  – Methods exist for calculating very complex graphs

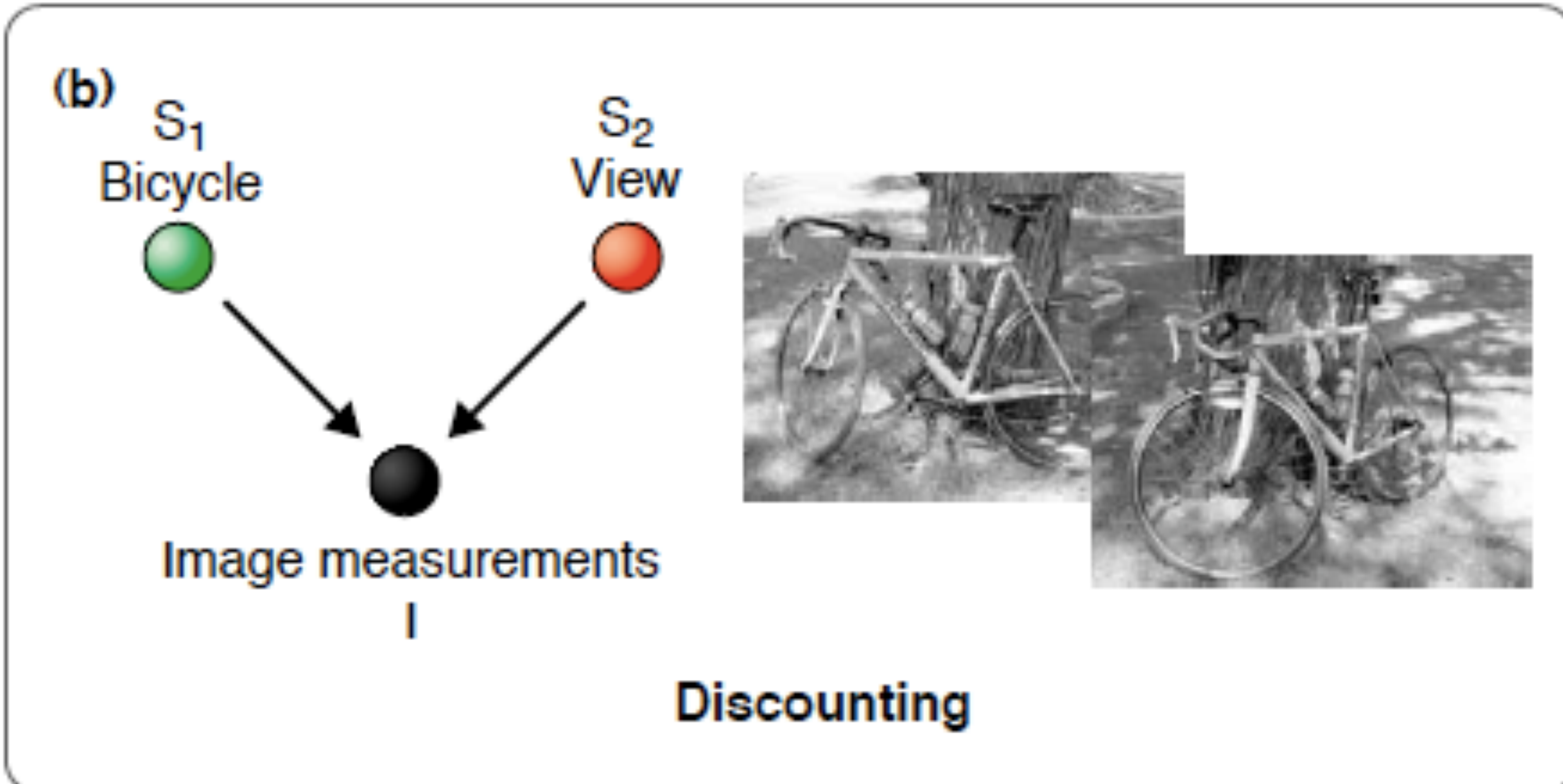
$$P(S \mid I) = \frac{P(I \mid S)P(S)}{P(I)}$$



Kersten & Yuille, 2003

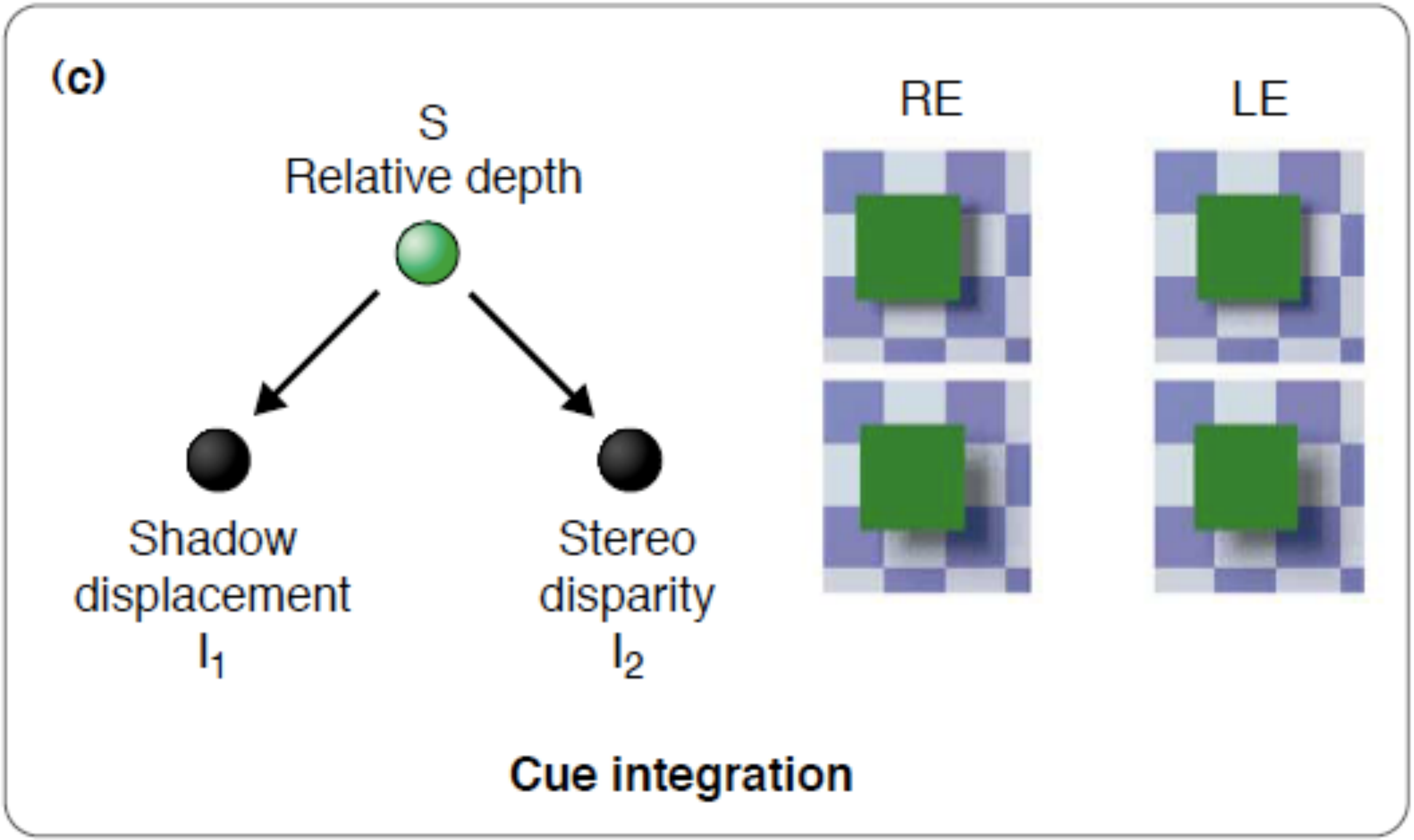
Graphical models

- Discounted cause is red for illustration



Kersten & Yuille, 2003

Cue integration



Kersten & Yuille, 2003

Cue integration

- Many cues to depth
  - Shadow is clearly a depth cue
  - Binocular disparity is a depth cue
- How do they interact?

Cue integration

- Strong fusion: Single underlying cause

$$P(S \mid I_1, I_2) = \frac{P(I_1, I_2 \mid S)P(S)}{P(I_1, I_2)}$$

- Conditional independence

$$P(S \mid I_1, I_2) = \frac{P(I_1, I_2 \mid S)P(S)}{P(I_1, I_2)} = \frac{P(I_1 \mid S)P(I_2 \mid S)P(S)}{P(I_1, I_2)}$$

Cue integration

- Gaussian likelihoods:

$$P(I_1 \mid S)P(I_2 \mid S) \propto$$

$$\exp\left(-\frac{(x-\mu_1)^2}{\sigma_1^2}\right)\exp\left(-\frac{(x-\mu_2)^2}{\sigma_2^2}\right) \propto$$

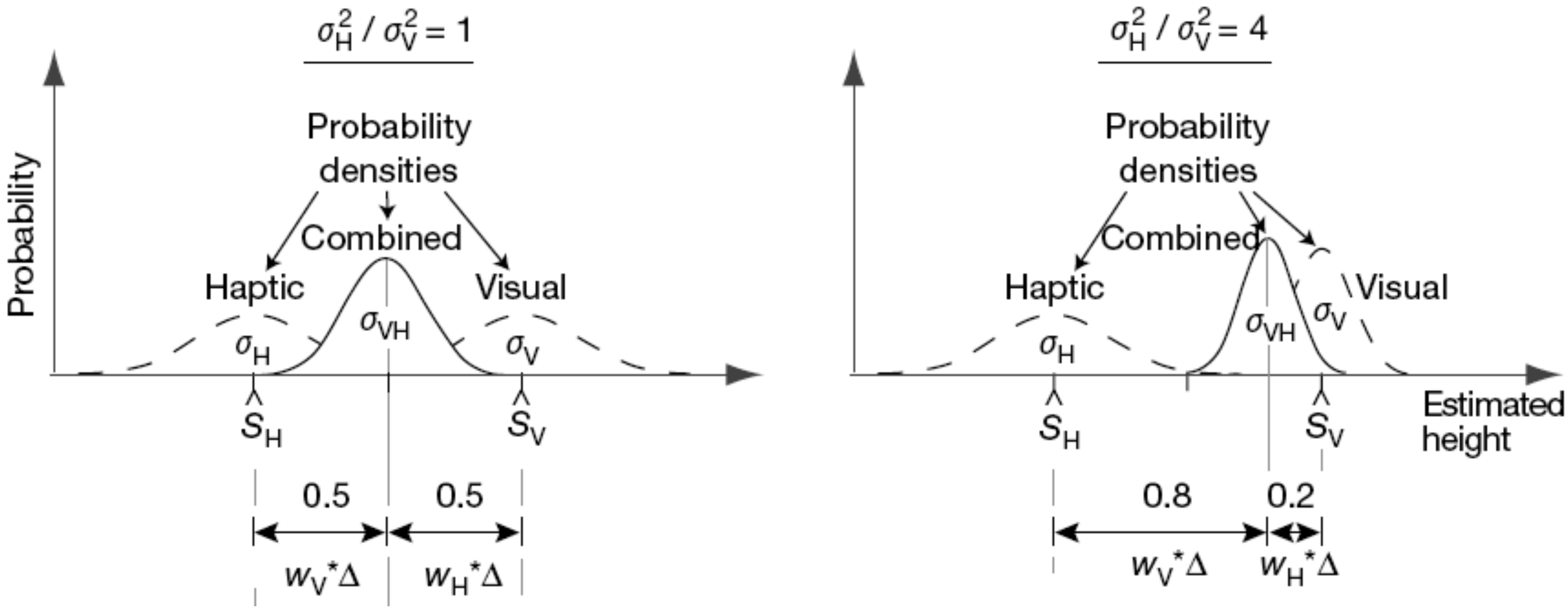
$$\exp\left(-\frac{(x-\mu_{fuse})^2}{\sigma_{fuse}^2}\right)$$

$$\mu_{fuse} = \frac{r_1}{r_1 + r_2}\mu_1 + \frac{r_2}{r_1 + r_2}\mu_2$$

$$r_i = \frac{1}{\sigma_i^2}$$

$$r_{fuse} = r_1 + r_2$$

Cue integration

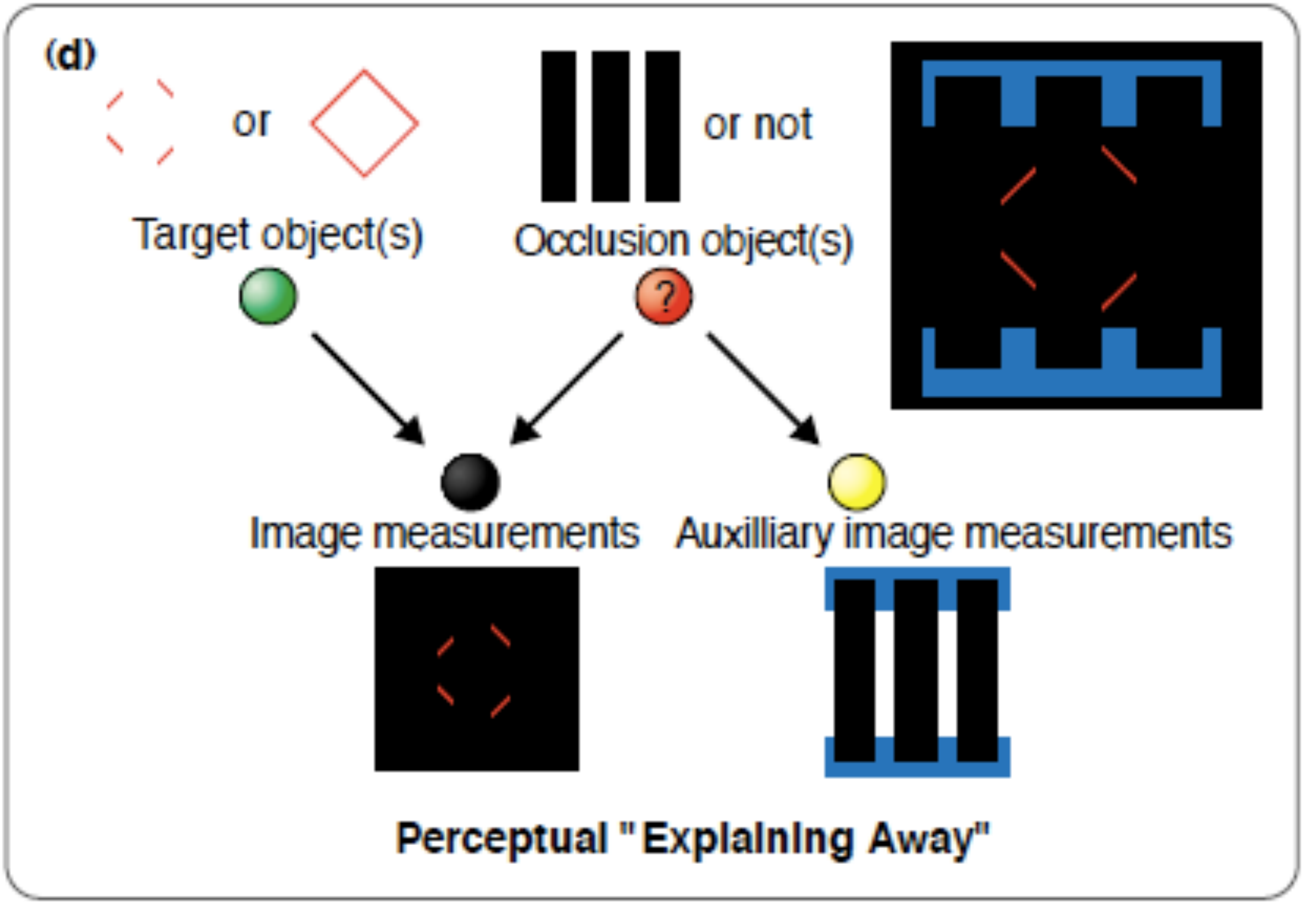


Ernst & Banks, 2002

Explaining away

- Example: Occluders
  - http://web.mit.edu/persci/demos/Motion&Form/demos/basic/basic-diamond.swf

Explaining away

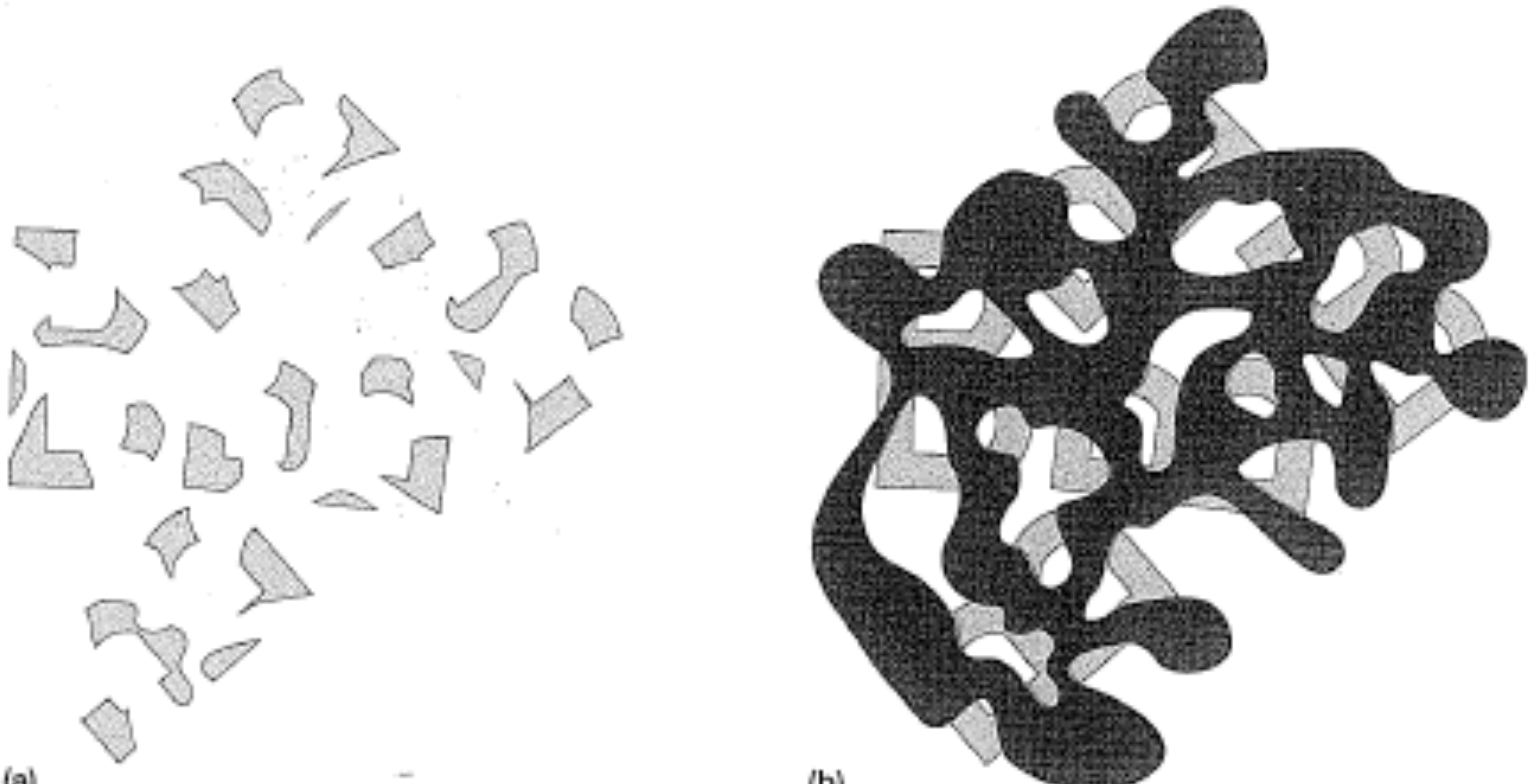


Kersten & Yuille, 2003

Explaining away

(a)

Explaining away



(a) (b)

Explaining away

- Acoustic analog:
  - Picket fence
  - Word restoration effect

Explaining away

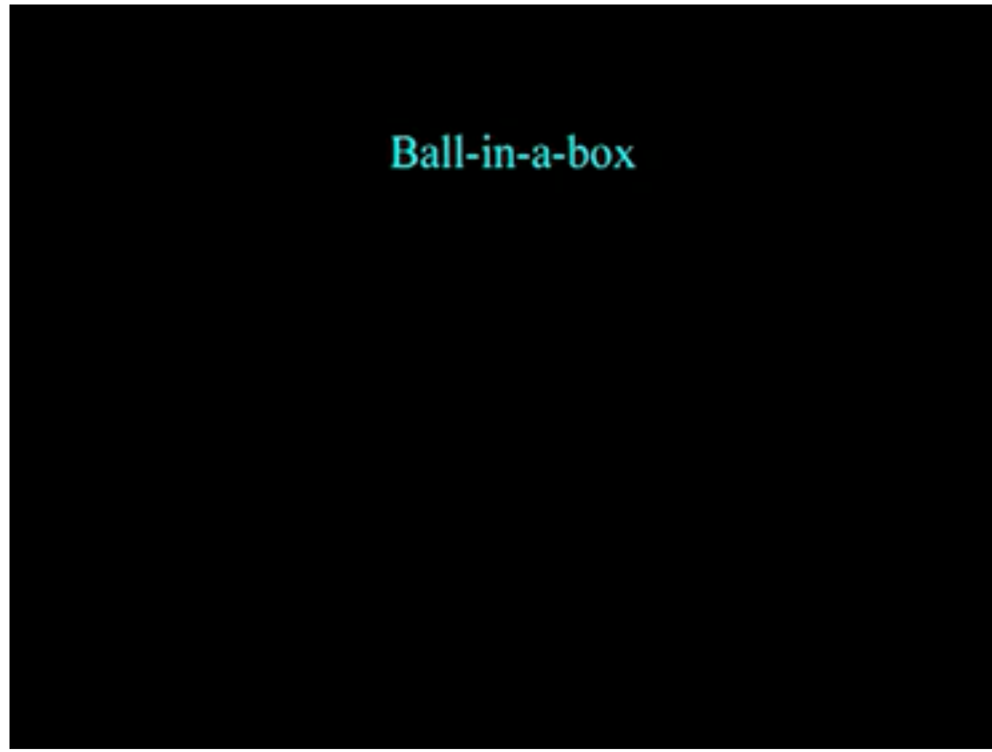
- Disambiguation through auxiliary evidence
- Reasoning analog
  - The lawn is wet
  - Did I leave the sprinkler on?
    - Or did it rain last night?
  - The neighbor’s lawn is wet
    - Decreases the probability that I left the sprinkler on

Explaining away

- Harry is late
  - Did he oversleep or are the busses on strike?
- Geoffrey is also late
  - Decreases the likelihood that Harry overslept

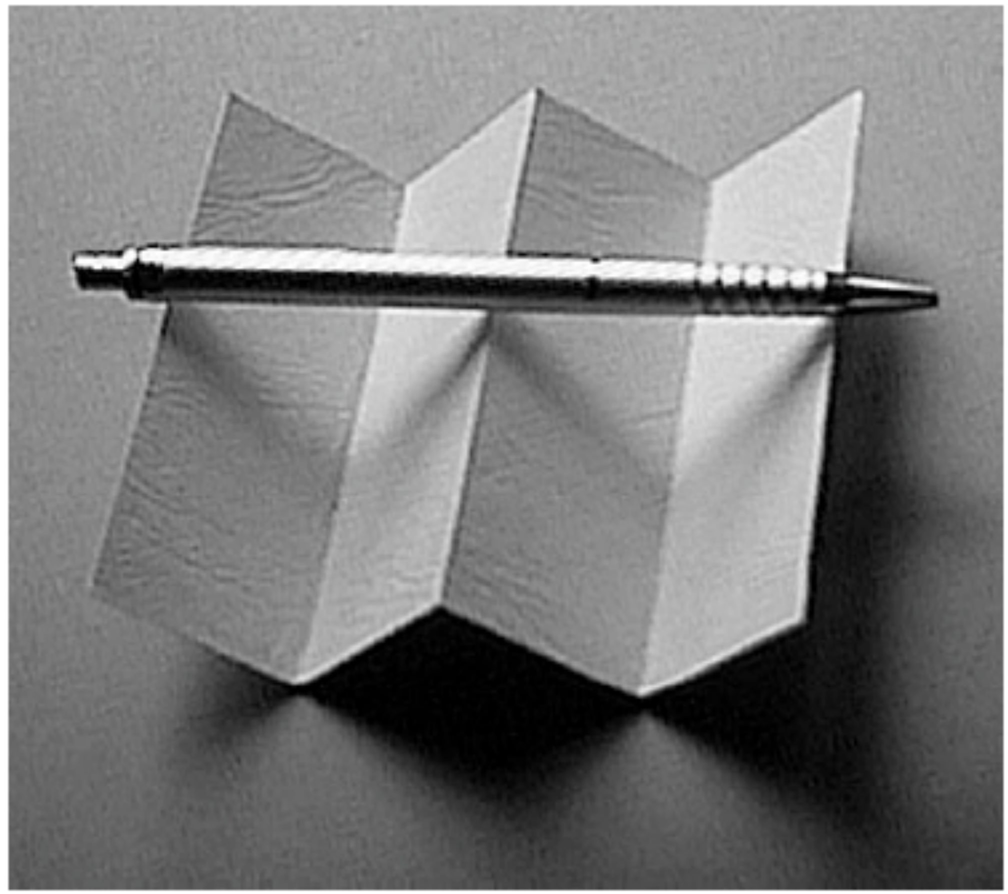
Explaining away

- Cast shadows disambiguates location (where?)



Explaining away

- Does not work for inferring shape (what?)



Neural implications

- Graphical structure suggest neural map
- Priors are analogous to top-down
- Likelihoods are analogous to bottom up
- Recurrent top-down and bottom up connections is explained by recursive Bayes
  – Today’s prior is yesterday’s posterior
- Problem remains: Encoding
  – How are probability distributions encoded?

Summary

- The Bayesian modeling framework is versatile
  - Discounting
    - Effects of task
    - Constancy
  - Cue (and prior) integration
    - Strong and weak fusion
  - Explaining away
    - Include auxiliary information