# Information and coding efficiency

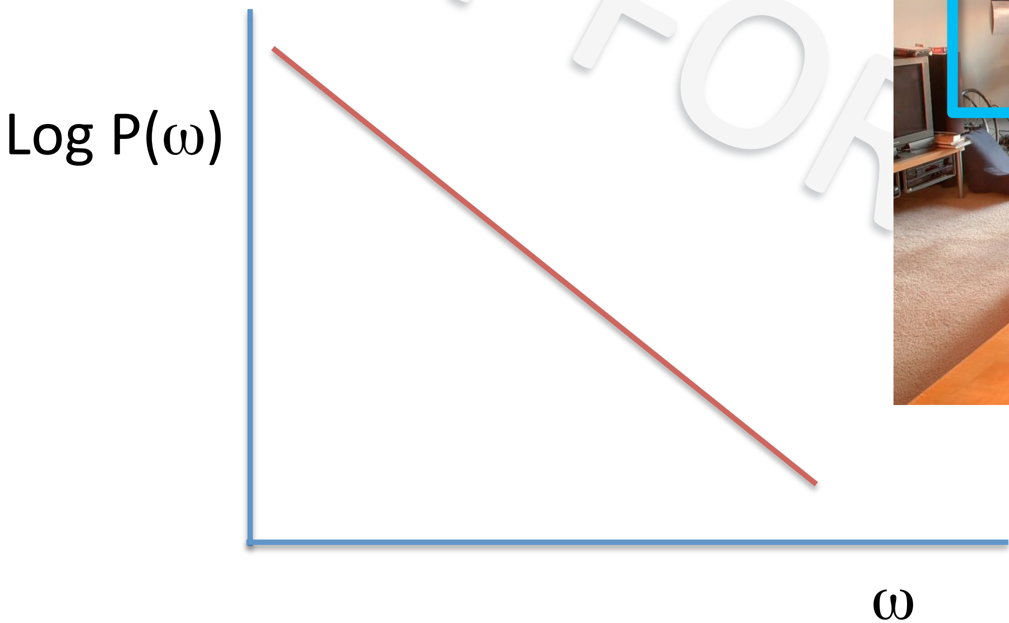
- What are the challenges posed by natural stimuli?
- What do information theoretic concepts suggest that neural systems should do?
- What principles seem to be at work in shaping the neural code?

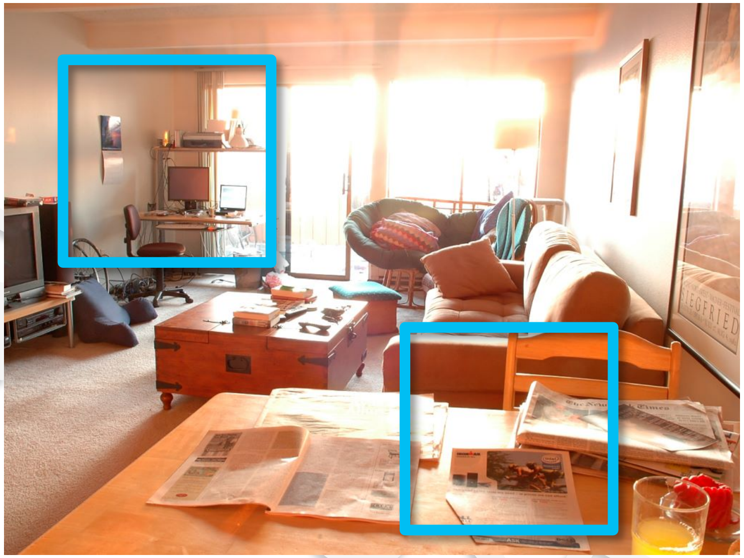
# Natural stimuli

我们的眼睛却可以很好地对焦，同时看清楚

1. Huge dynamic range: variations over many orders of magnitude

2. Power law scaling

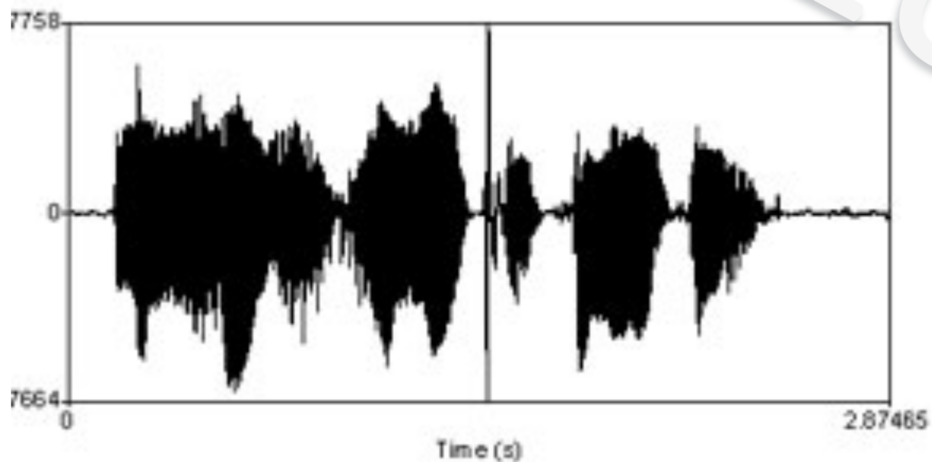


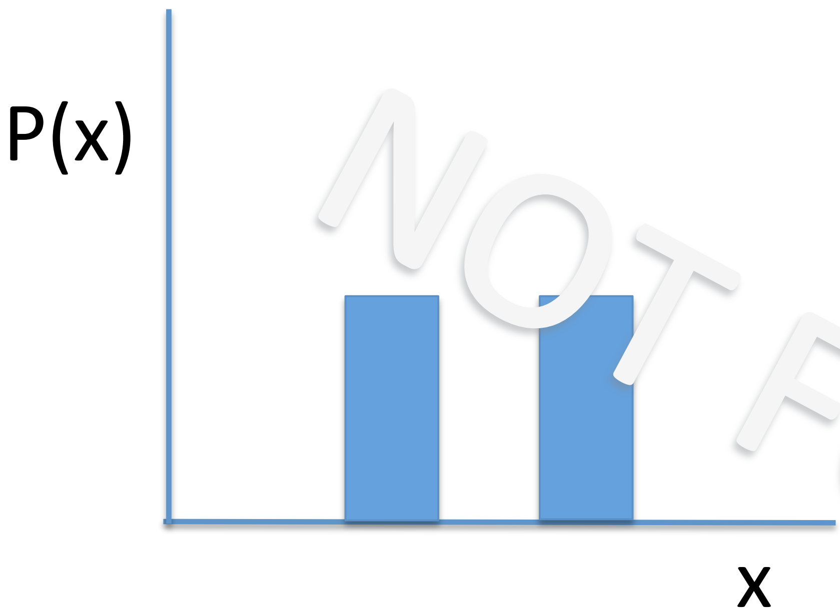
# Natural stimuli

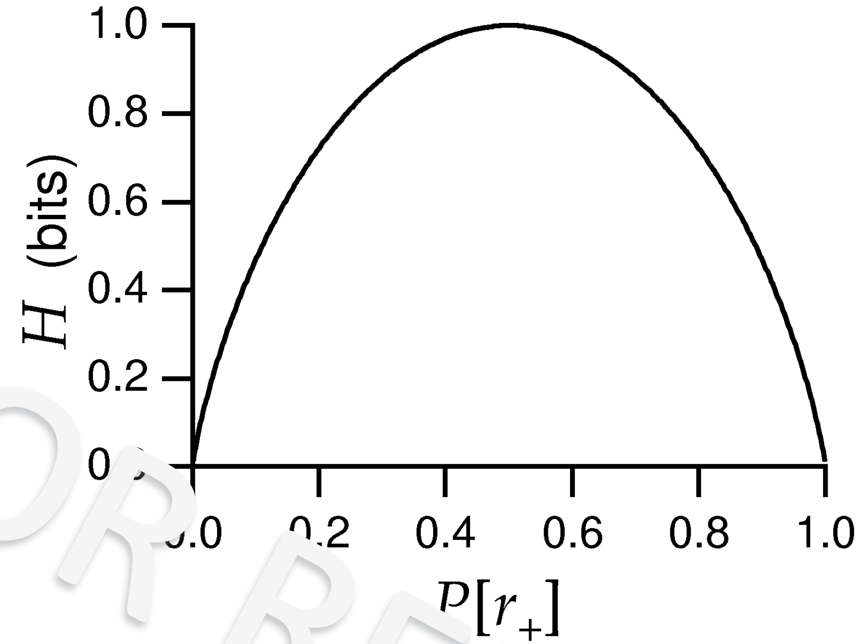
1. Huge dynamic range: variations over many orders of magnitude

2. Structure at many scales



# What makes a good code?






Entropy = $- \Sigma\, p_i \log_2 p_i$

Maximize information:

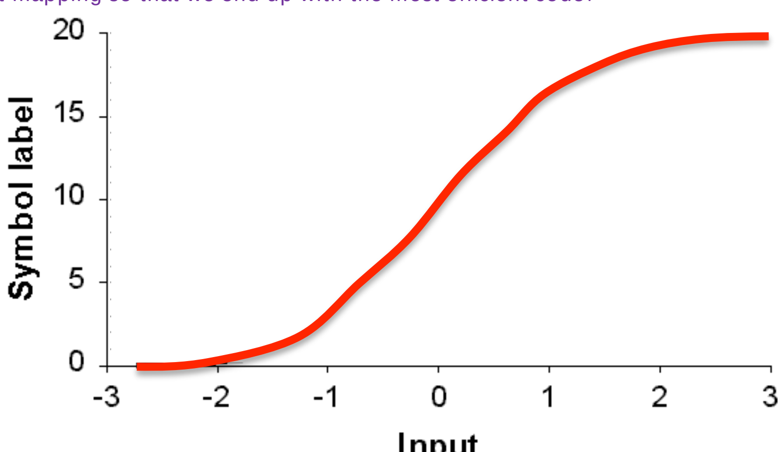
$I(R,S) = H[R] - \langle H[R|s] \rangle_s$

可控 不可控

# Efficient coding

In order to have maximum entropy output, a good encoder should match its outputs to the distribution of its inputs


We have some limited range of possible symbols at our disposal, say zero to 20 hertz.
How should we organize that mapping so that we end up with the most efficient code?




Maximize information:

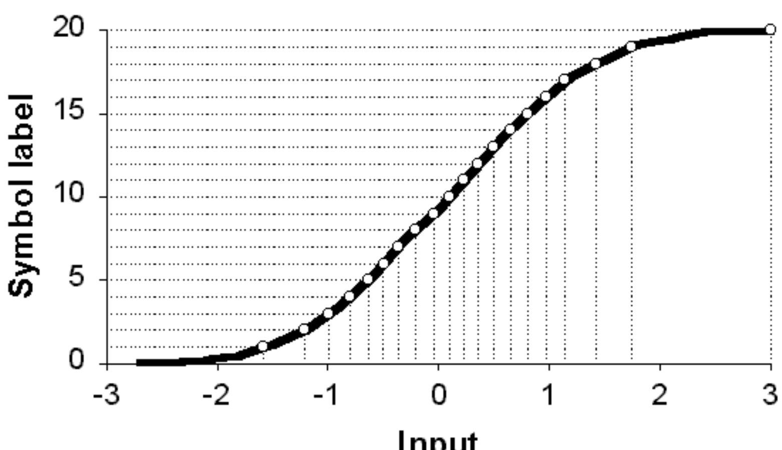
That is, by using all of our symbols about equally often

$$I(R,S) = H[R] - \langle H[R|s] \rangle_s$$

Barlow, Attneave 1954

# Efficient coding

In order to encode stimuli effectively, an encoder should match its outputs to the statistical distribution of the inputs


The input/output function should be determined by the distribution of natural inputs

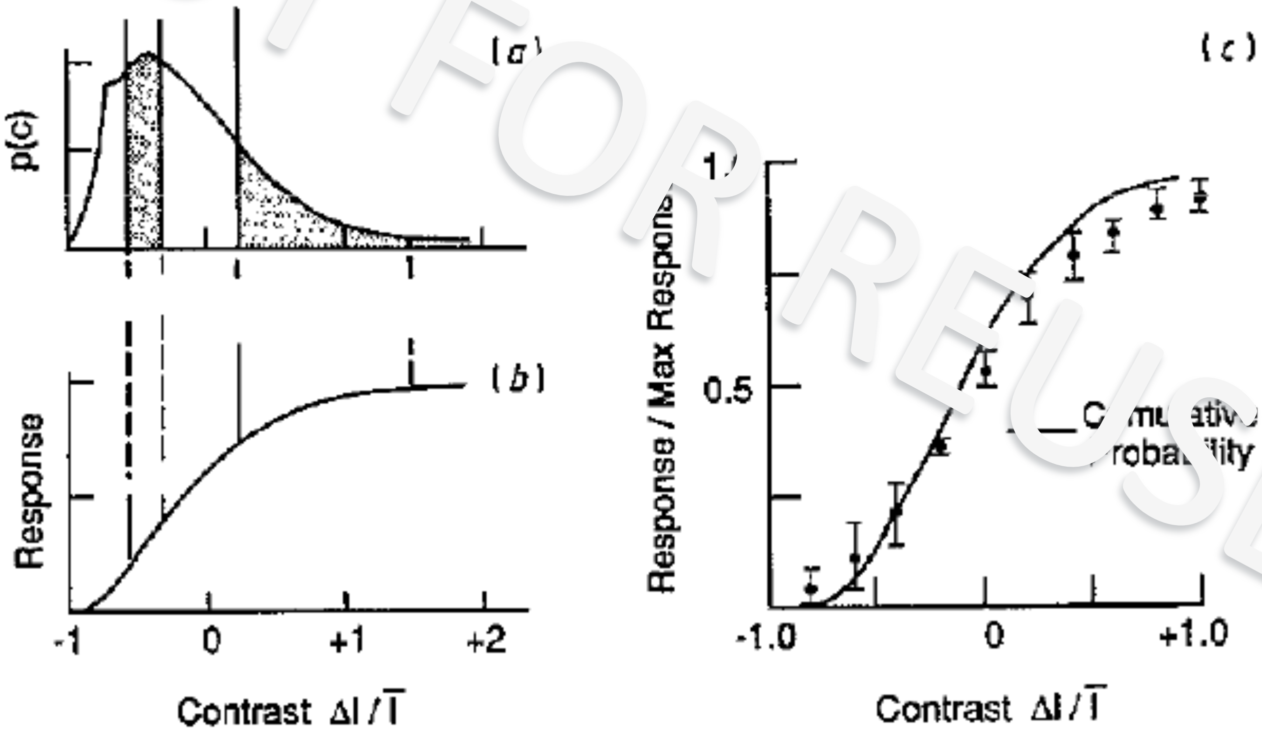
Optimizes **mutual information** between output and input

Barlow, Attneave 1954

# Fly visual system

$$P(r)dr = P(s)ds$$

$$r = g(s) = \frac{1}{\alpha}\int_{-1}^{s} ds'\, P(s').$$




Laughlin, '81

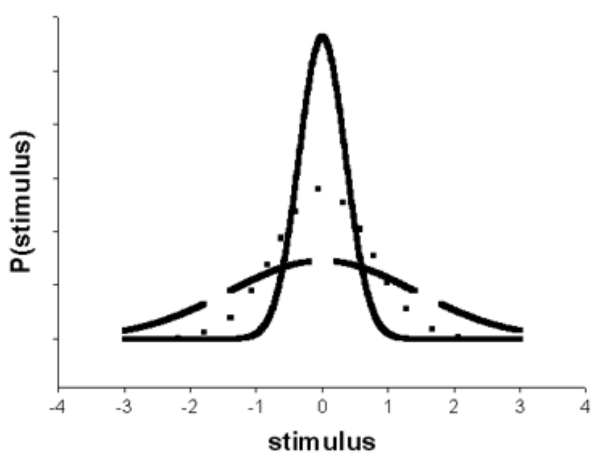
# Variation in time

Contrast varies hugely in time.

Should a neural system optimize over evolutionary time or locally?

上面的对比度很大，stimulus的分布就很宽，标准差很大

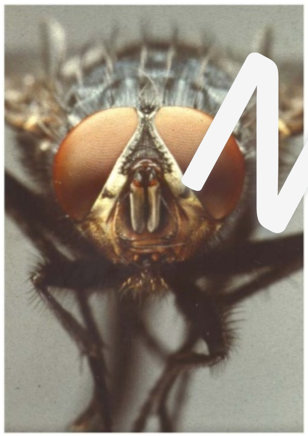
下面的对比度很小，stimulus的分布就很窄，标准差也很小

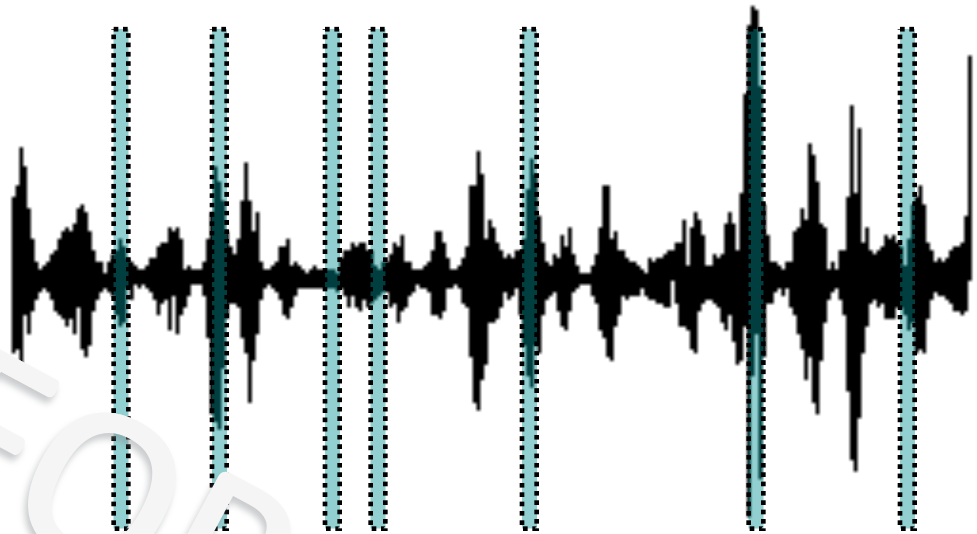



What our code should do is take the widths of these distributions into account in setting up a local. Input, output curve, that accommodates this structure of the, currently measured statistics of the input

# Time-varying stimulus representation

input: white noise × time vary envelope(σ(t))





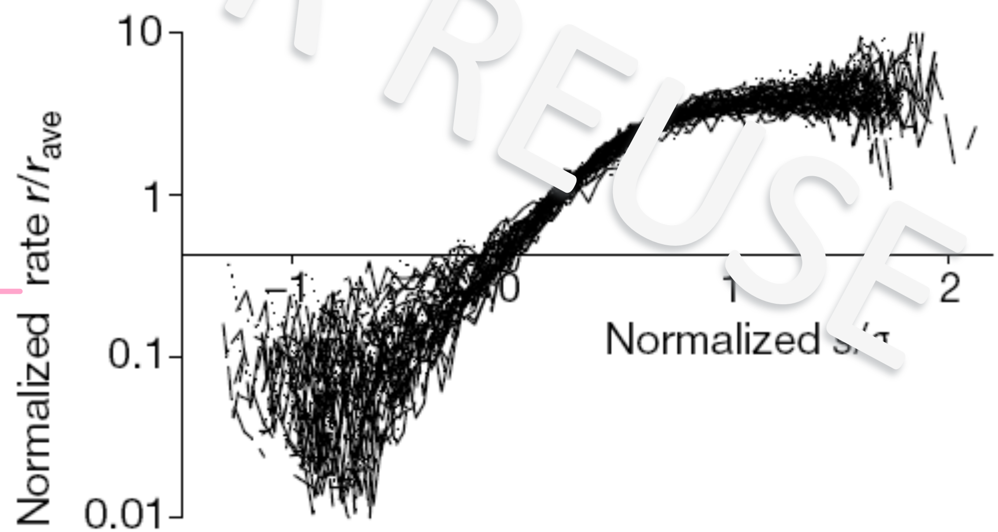
For fly neuron H1, determine the input/output relations throughout the stimulus presentation

σ不变，改变白噪音，得到input-output function

当刺激的幅度比较大的时候，input-output function就比较平坦；当刺激幅度比较小的时候，就比较陡峭。

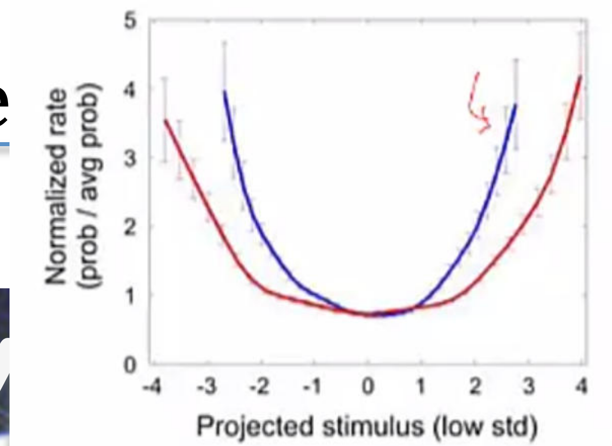
如果我们把input给标准化了，就是把envelope去掉，发现input-output function变得差不多。

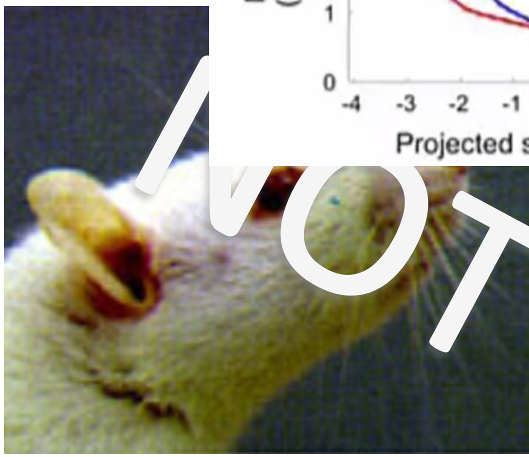
What that says is that the code has the freedom to stretch its input access such that it's accommodating these variations in the overall scale of the stimulus.

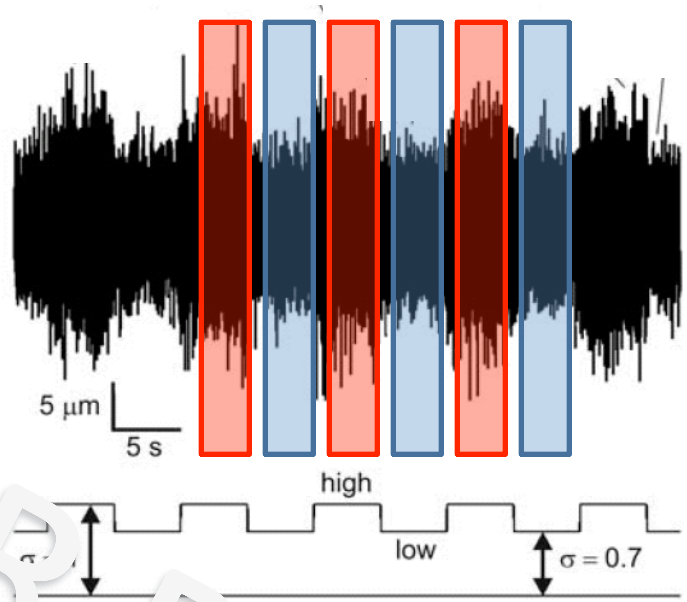


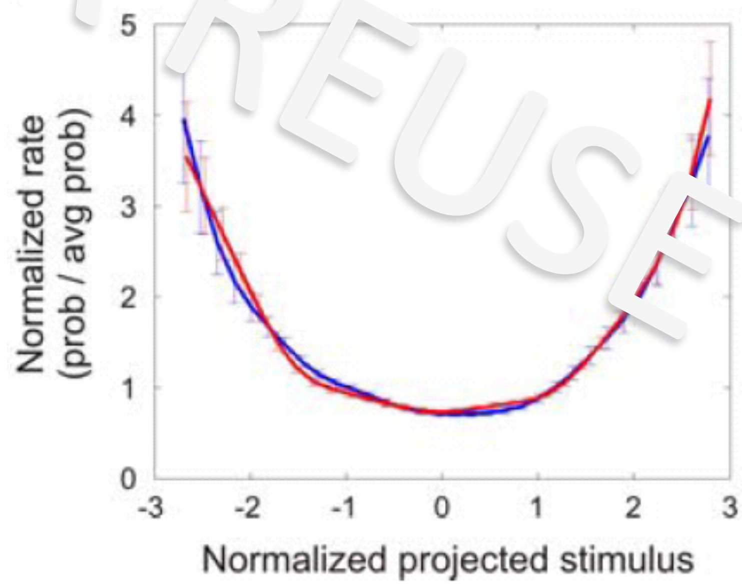

Fairhall, Lewen, de Ruyter and Bialek (2001)

# Barrel corte





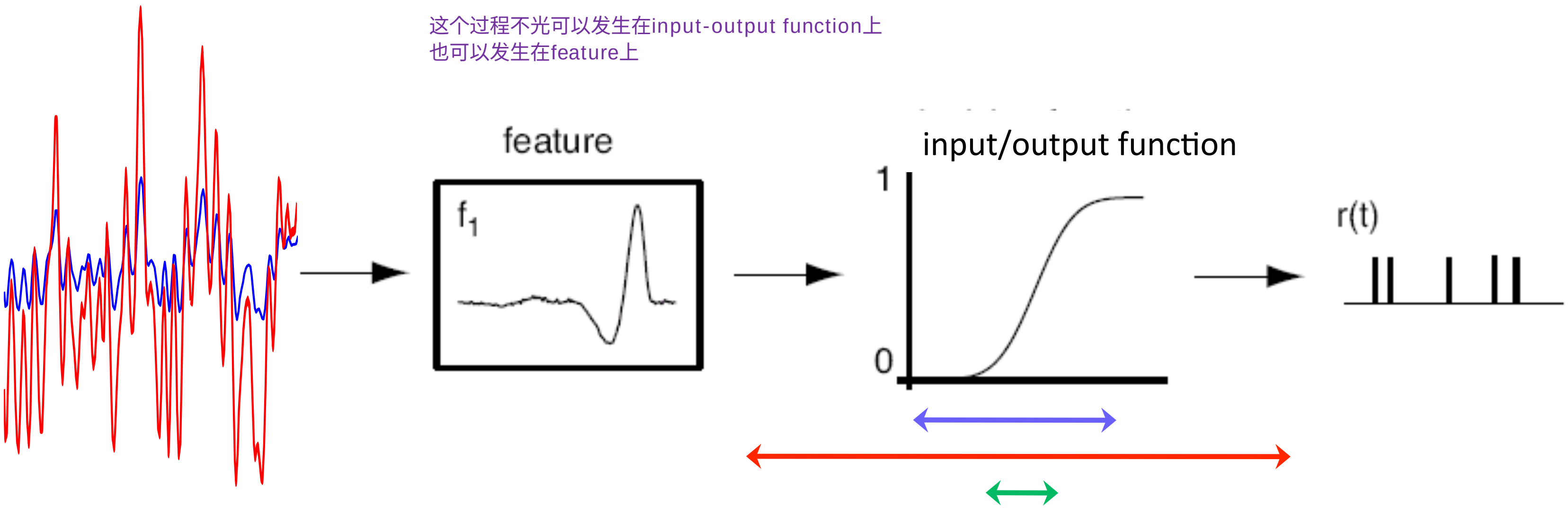




Extracellular *in vivo* recordings of responses to whisker motion in rat S1 barrel cortex in the anesthetized rat

M. Maravall et al., (2007)

# Adaptive representation of information

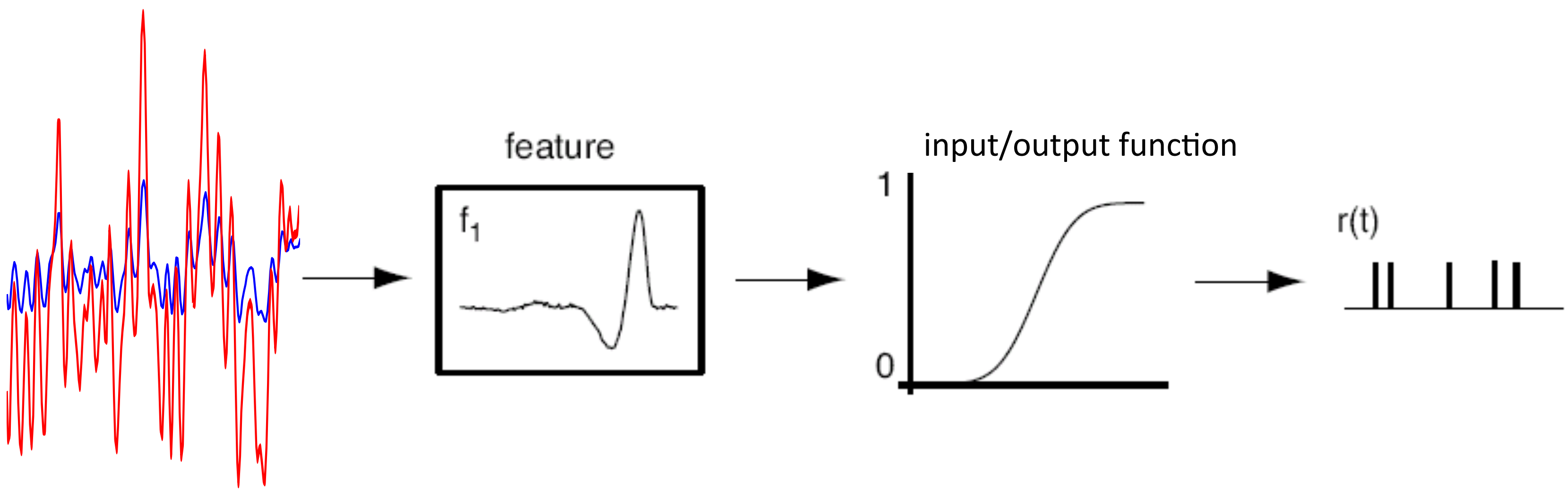
这个过程不光可以发生在input-output function上
也可以发生在feature上




As one changes the characteristics of $s$(t), changes can occur in the *input/output function* and in the encoded *feature.*

Barlow '50s, Laughlin '81, Shapley et al, '70s, Atick '91, Brenner '00

# Feature adaptation



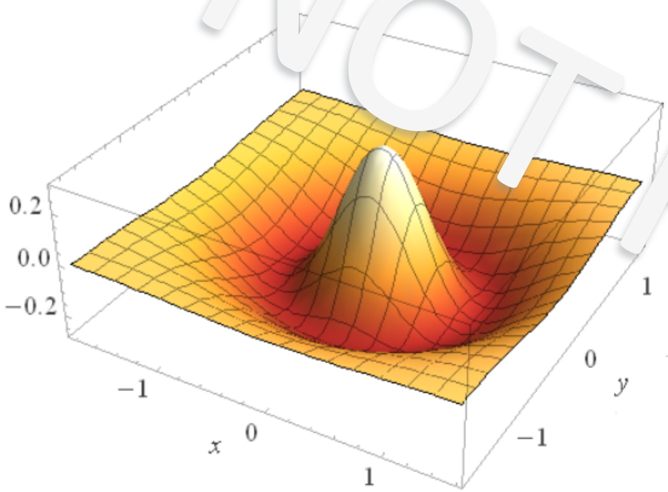

Barlow ’50s, Laughlin ‘81, Shapley et al, ‘70s, Atick ‘91, Brenner ‘00

# Feature adaptation

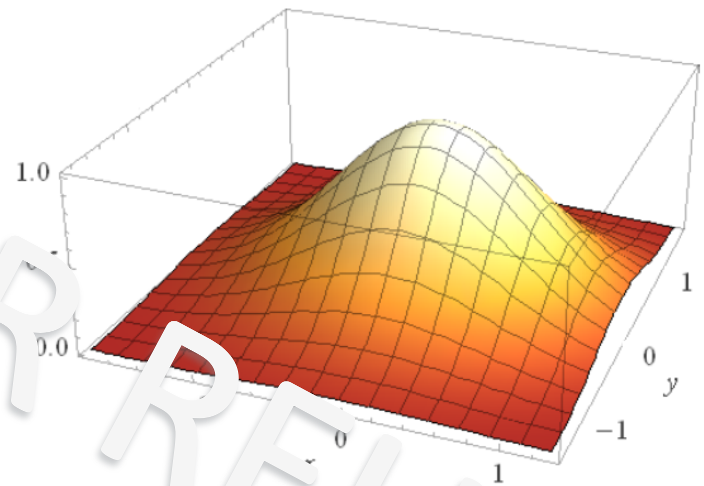
If we consider that the retina imposes a linear transfer function, or a filter on its inputs, what's the shape of that filter that maximizes information transmission through the retina?

The solution turns out to depend on two things. The powers spectrum of natural images and the signal-to-noise ratio.






Predicted receptive field at high light level




Predicted receptive field at low light level

At high light levels, or high signal to noise, one would predict a filter shape in the left. This acts like a differentiator, looking for edges of the stimulus, but at low light levels, the predicted optimal filter is integrating, and simply averages its inputs to reduce noise. And indeed in retinal receptive fields it's seen that the surround becomes weaker at low-light levels and the center broader which qualitatively matches these predictions.

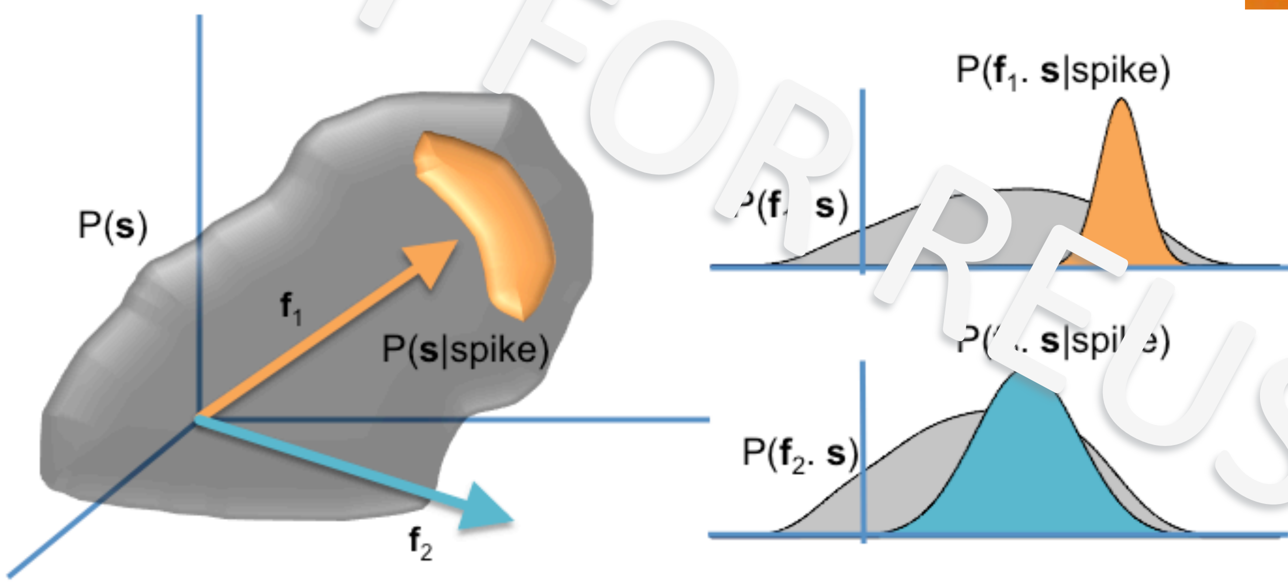
Atick and Redlich ('90), Atick ('92)

# Feature adaptation

Choose filter in order to maximize $D_{KL}$ between spike-conditional and prior distributions.

Equivalent to maximizing information that the spike provides about the stimulus.

Return of the maximally informative dimension!




Sharpee, Rust and Bialek, Neural Computation, 2004

# Redundancy reduction

如果只为了效率以及能有更多的信息，每个神经元应该尽可能独立。

Population code: $P(R_1,R_2)$

$$H[R_1,R_2] \leq H[R_1] + H[R_2]$$

However.. correlations can be good.

- Error correction and robust coding
- Correlations can help discrimination

Indeed, neurons in the retina are observed to be redundant (Berry, Chichilnisky)

# Representing natural scenes sparsely

Neuron populations should be as sparse as possible. That is that their coding properties should be organized so that as few neurons as possible are firing at any time.

系数 basis function noise

$$I(\vec{x}) = \sum_i a_i\, \phi_i(\vec{x}) + \epsilon(\vec{x})$$

Fourier basis is guaranteed to be able to represent any image, but coding with such a basis is not sparse. Because the power spectrum is broad.

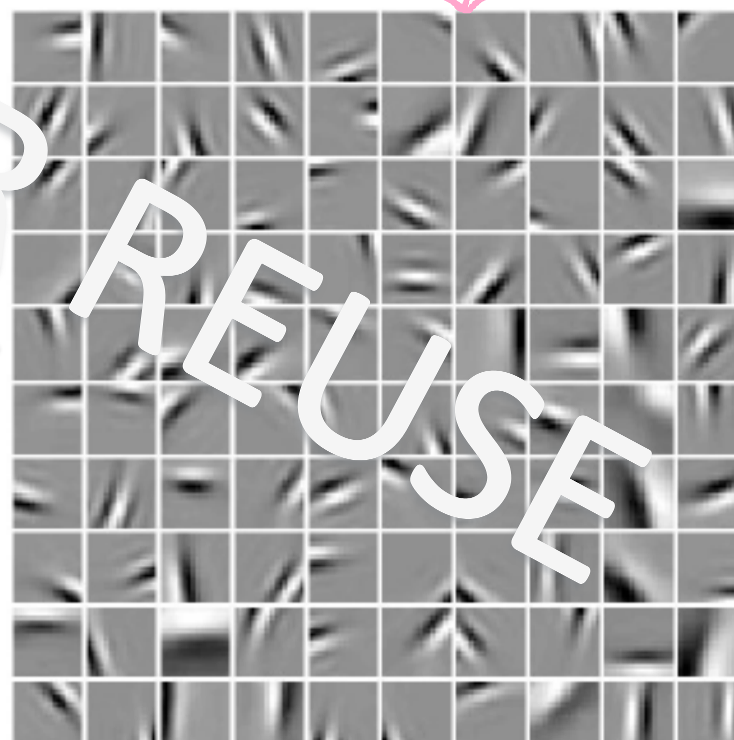
reconstruction error

选择需要尽量少的系数的basis function

$$E = \sum_{\vec{x}} \left[ I(\vec{x}) - \sum_i a_i\, \phi_i(\vec{x}) \right]^2 + \lambda \sum_i C(a_i)$$

用最少的neuron表达input

Sparse code: When we view an image using neuronal receptive fields that look like this, this excites on average a minimal number of neurons



Olshausen; Olshausen and Field (1996), Bell and Sejnowski (1995)

# Coding principles

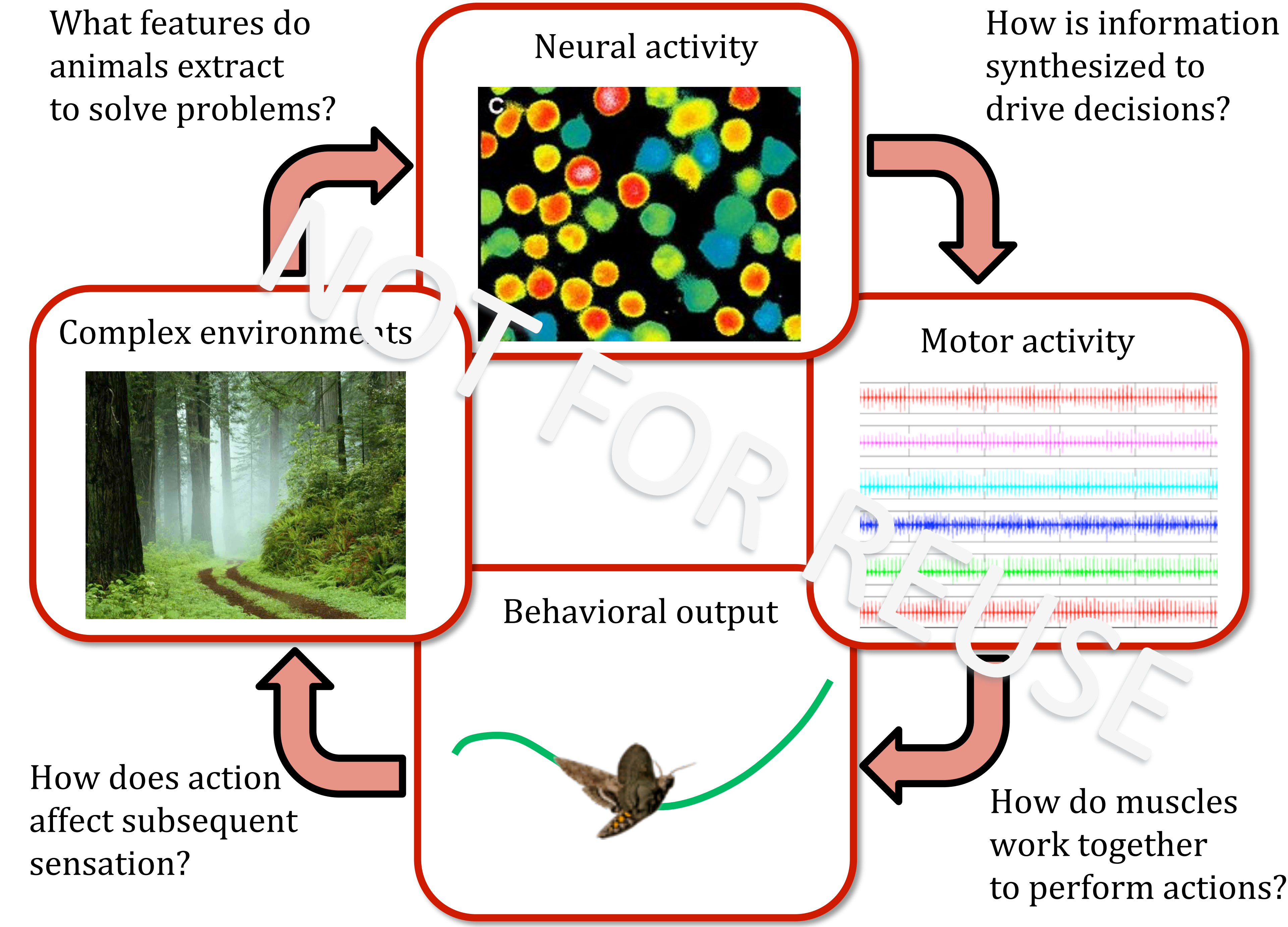
- Coding efficiency
- Adaptation to stimulus statistics
- Sparseness

# And so to the end of coding…

Classic and state of the art methods:

- Models for predicting how stimuli are coded in spikes
- Models for decoding stimuli from neural responses
- Information theory and how it is used to evaluate coding schemes
- A very quick glance at how coding strategies might be shaped by the statistics of natural inputs

# What have we missed?


What features do animals extract to solve problems?
Neural activity
c
How is information synthesized to drive decisions?
Complex environments
Motor activity
Behavioral output
How does action affect subsequent sensation?
How do muscles work together to perform actions?

# Next week

A brief introduction to the biophysics of coding

- Neuronal excitability
- Simplified models that capture neuronal firing