Neurophysiology and Information

Christopher Fiorillo

BiS 527, Spring 2010

042 350 4326, fiorillo@kaist.ac.kr

Part 6: Efficient Coding of Sensory Information

Reading:

"Spikes: Exploring the Neural Code" by Rieke et al., 1997. Chapters 2 and 3.

"Natural Image Statistics and Neural Representation" by Simoncelli and Olshausen, Annual Reviews of Neuroscience, 2001.

"Efficiency and Ambiguity in an Adaptive Neural Code" by Fairhall et al., Nature 2001.

Information Theory

- Information theory began with Claude Shannon publication in 1948 of "A mathematical theory of communication."
- He defined information as a reduction in entropy. This was his most important accomplishment, in my opinion.
- An efficient communication channel
  - He was working for a telephone company, and his interest was in finding an efficient code for transmitting information through a communication channel, from sender to receiver.
  - The communication channel has a limited capacity, which is equal to its entropy.
  - If the channel is transmitting information between sender and receiver, what code should be used to maximize the mutual information (shared information)?

Information Theory

- Information theory was developed from the third-person perspective of the engineer. It uses primarily information derived directly from measurable frequency distributions.
- For Shannon's purposes, this was the natural and appropriate perspective.
- Imagine that two people are speaking on a telephone.
  - They both know the English language very well, and the engineer who designed the telephone system also knows English very well.
  - Thus, the probabilities of words are approximately the same for all three people.
    - Actually this is not very accurate, since the people having the conversation know about its unique content, and can use that information to make better predictions about words. Shannon calculated the entropy of the English language, but there is a lot of information available to people who are having a conversation that Shannon did not utilize.

Information Theory

- Information theory is easier to understand if we imagine situations in which each possible state (such as a word) is equally probable.
- Imagine that there are X words in a language. Thus the receiver can guess that the sender will speak one of these words at a particular moment in time. In this case the prior entropy of the receiver's prediction will be proportional to log (X). If the receiver correctly perceives the one word, then his posterior entropy will be zero, and the mutual information between sender and receiver will be proportional to log(X). This is perfect communication.
- But if the communication channel can only convey 1 bit per second (1 bit distinguishes two possible sets of words), then it will take at least several seconds for the channel to specify one word. The actual amount of time will depend on the code that is used.

Information Theory

- In general, not all possible states are equally probable.
- Entropy = -sum(p(log(p))
- Mutual information = entropy (X) - entropy (X|Y)
- Mutual information is the amount of information that is shared (between sender (X) and receiver (Y))
  - It can also be thought of as the amount of information that is transmitted.
- The goal of efficient coding is therefore to find a code that maximizes the mutual information, given limited channel capacity and whatever knowledge is available about the probabilities of "messages"
  - In general, this needs to be done in the presence of "noise"
-

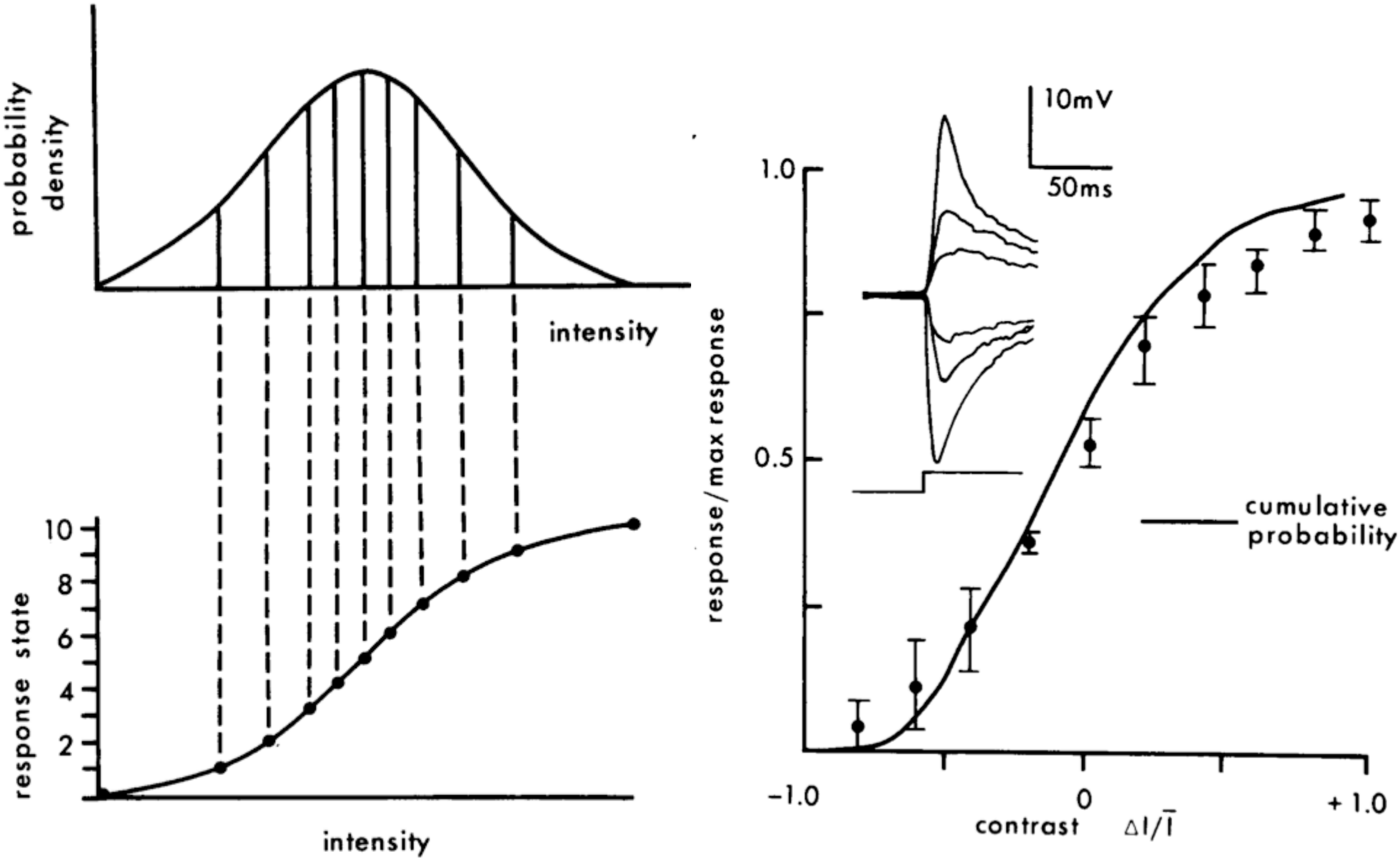
Efficient Coding in Neurons

- Shortly after Shannon published his work on information theory, neuroscientists began to apply his ideas
- Barlow (1962) was the first to do this with respect to the physiology of single neurons, particularly in the retina
- The idea of efficient coding has been applied to sensory systems in general.
- A neuron has a limited number of discriminable outputs, and thus it has a finite coding (information) capacity.
- The idea is that a neuron should function as an efficient communication channel.
- Using Shannon's terms, the sensory environment is the "sender," and the "receiver" is the downstream neuron
- A lower sensory neuron should send as much information as possible to higher sensory neurons.
- Barlow proposed that sensory neurons should maximize mutual information by avoiding redundancy

Redundancy and Predictability

- Sensory stimuli (like light intensity) are correlated across space and time
  - Thus they are said to be "redundant"
  - Redundant means predictable
- Redundancy (predictability) reduces the information content of the sensory stimuli
- To be efficient, the neuron should transform its input (corresponding to the sensory signal) to remove the redundancy. The neuron should not signal the predictable component of the sensory signal.

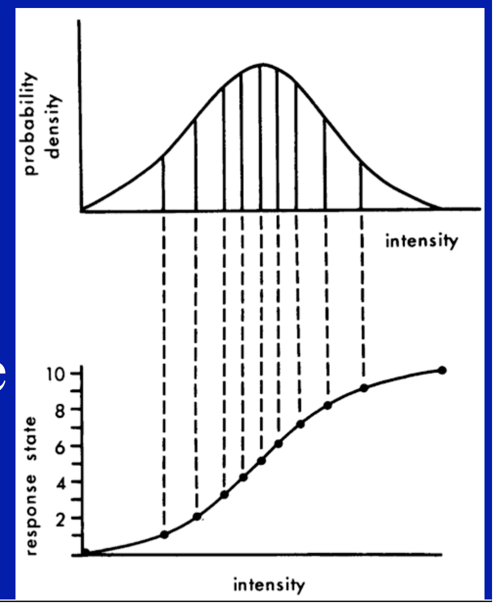
Efficient Coding in a neuron of the fly eye



- A neurons I-O curve should be matched to the frequency distribution of its inputs
- Laughlin measured naturally occurring contrasts, and showed that a neuron's I-O curve actually does match the cumulative frequency distribution of natural contrasts
- The neuron discriminates well amongst the commonly occurring intensities, but it sacrifices its ability to discriminate between intensities that are not common. This is efficient use of limited capacity.

An Efficient Coding Strategy for a Gaussian Distribution of Inputs

- If an input has a Gaussian frequency distribution, then the mean and standard deviation are constant over time (invariant).
- Thus they are predictable and redundant, and should be removed from the output.
- This can be done by signaling stimulus intensity as the difference from the mean in units of standard deviation
  - This is called a Z-score in statistics
- Neurons can adapt quickly, over just a few seconds, to changes in mean and variance
  - This has been shown in motion-sensitive neurons of the fly, and in dopamine neurons of monkey



The Neuron’s Point of View

- The efficient coding literature implies that information can be objectively quantified without regard to perspective.
  - As we have discussed, this view is completely wrong.
- Whether a neuron’s output is an efficient representation of its input depends on what prior information the neuron has.
- If the neuron has no prior information, then it is perfectly efficient, from the neuron’s perspective, for output to equal input. In this case, the stimulus is not predictable for the neuron.
- From our third-person perspective, in order for the neuron to have a perfectly efficient code, it would need to have exactly the same prior information that we have
- When we observe that a neuron adapts to the mean and variance of a stimulus distribution, that means that the neuron, or some component of the nervous system, has information about the mean and variance.

Conclusions

- The efficient coding hypothesis has been the most successful computational theory in neuroscience, in my opinion
  - It has explained many aspects of sensory systems
- Limitations of the efficient coding hypothesis:
  - "Efficiency" depends on the knowledge of the person observing the system
    - 'Intermediate' level sensory neurons appear to be efficient because their prior information is similar to our prior information.
    - The "first" sensory neurons (e.g. photoreceptors) appear rather inefficient because they have less prior information about light intensity than scientists
    - The efficiency of higher sensory neurons cannot be judged because they respond to parts of the world for which we do not know about the statistics
  - Efficient communication is not the primary function of the nervous system.
    - But efficiency requires information, and thus prediction. Prediction is the primary function of the system from the first-person perspective.
  - Information theory does not tell us anything about which information is the most important