PROBABILITISTIC MODELS FOR BIOLOGICAL OBSERVATIONS

One of the common major goals of the work described in the previous section was the derivation of simple models to help understand complex biological processes. As these models evolve, they not only can help us understand, but also they suggest aspects that experimental methods alone may not. In part, this is because the mathematical model allows for greater control of the (simulated) environmental conditions. This control allows the researcher to, for example, identify stimulus-response patterns in the mathematical model whose presence, if verified experimentally, can reveal important insight about the intracellular mechanisms.

At the workshop, John Rinzel of New York University explained how he had used a system of differential equations and dynamical systems theory to model the neural signaling network that seems to control the onset of sleep. Rinzel's formulation sheds light on the intrinsic mechanisms of nerve cells, such as repetitive firing and bursting oscillations of individual cells, and the models were able to successfully mimic the patterns exhibited experimentally. More detail may be accessed through his web page at http://www.cns.nyu.edu/corefaculty/Rinzel.html.

In another approach, based on point processes and signal analysis techniques, Don Johnson of Rice University formulated a model for the neural processing of information. When a neuron receives an input (an increase in voltage) on one of its dendrites, a spike wave—a brief, isolated pulse having a characteristic waveform—is produced and travels down the axons to the pre-synaptic terminals (see Figure 7). The spike waves occur randomly in time, and sensory information in the nervous system is embedded in the timing of the spike waves. These spikes are usually modeled as point processes; however, these point processes have a dependence structure and, because of the presence of a stimulus, are non-stationary. Thus, non-Gaussian signal processing techniques are needed to analyze data recorded from sensory neurons to determine what aspects of the stimulus are being emphasized and how emphatic that representation might be.

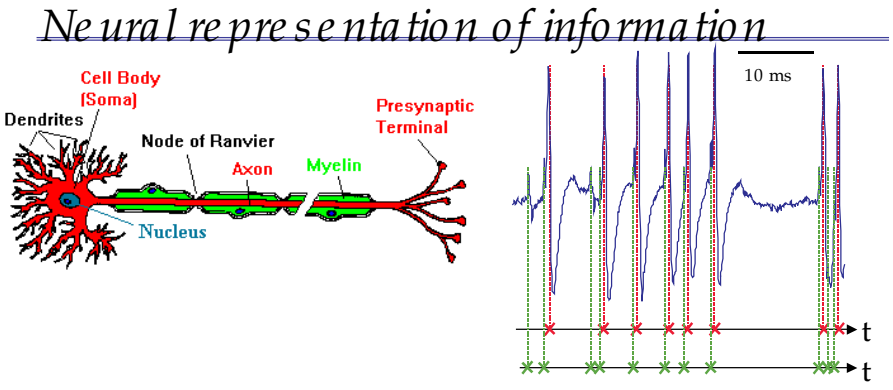
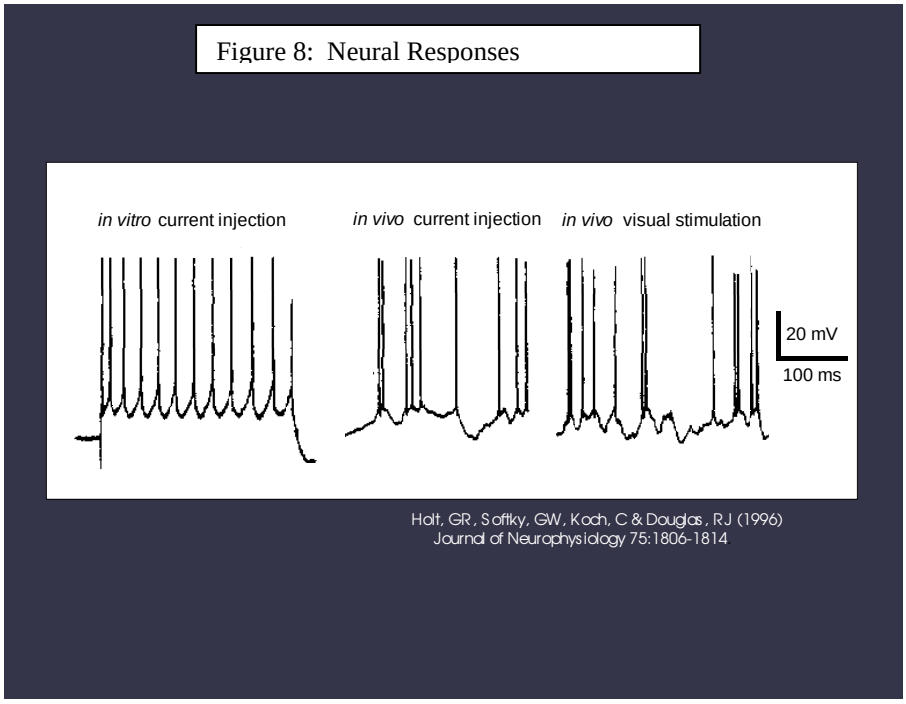


Figure 7: Neural Representation of Information

Johnson developed the necessary signal processing techniques and applied them to the neuron spike train. Details may be found in Johnson, *et al.* (2000), and also

http://www.ece.rice.edu/~dhj/#auditory. This theory can be extended to an ensemble of neurons receiving the same input, and under some mild assumptions the information can be measured with increasing precision as the ensemble size increases.

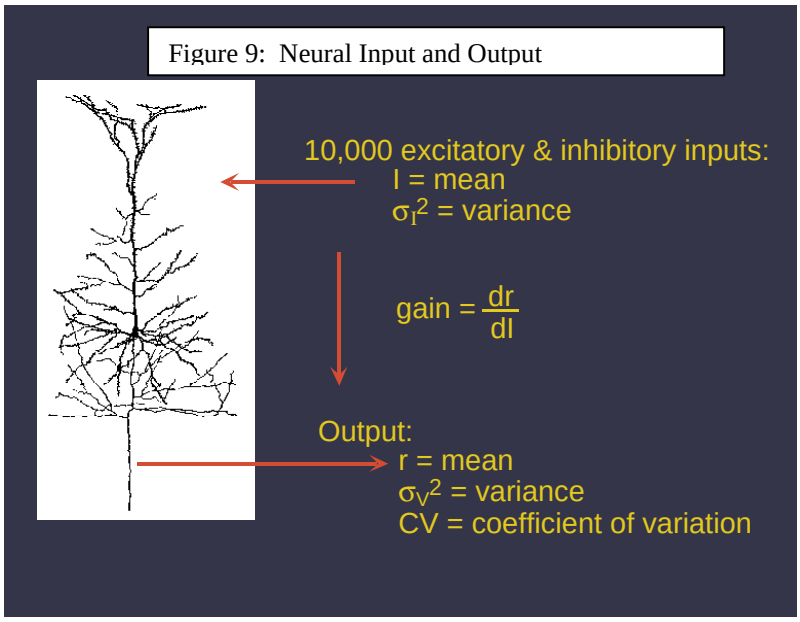


Larry Abbott of Brandeis University also explored the characteristics of neuron signals. At the workshop, he presented research on the effect of noise as an excitatory input to obtain a neural response, and his methods took advantage of the difference between *in vivo* measurements and *in vitro* measurements. His work counters one of the most widely spread misconceptions, that conductance alone changes the neural firing rate. Instead, a combination of conductance and noise controls the rate. As Figure 8 shows, although a constant current produces a regular spike train *in vitro*, this does not happen *in vivo*, where there is variance in the response, thus more noise in the signal.

It is of great interest to study the input and output relations in a single neuron, which has more than 10,000 excitatory and inhibitory inputs. Let I denote the mean input current, which measures the difference between activation and inhibitory status, and let σ_I^2 be the input variance. For an output with a mean firing rate of r, neuroscientists typically study the output's σ_v^2 and coefficient of variation CV. Abbott also studies the gain dr/dI. See Figure 9. The standard view is that

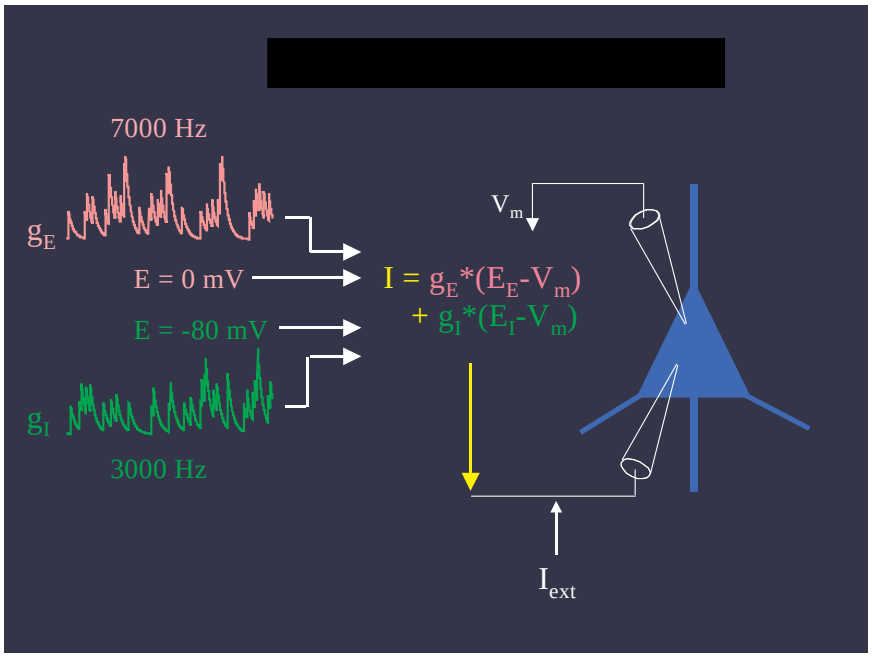
- the mean input I controls the mean output r, and
- the variance of the input affects σ_v^2 and CV.

Abbott disputes the second part and concludes that the noise channel also carries information about the firing rate r. To examine this dispute, Abbott carried out *in vitro* and *in vivo* current injection experiments.



In the first experiment, an RC circuit receiving constant current was studied. Such a circuit can be represented with a set of linear equations that can be solved analytically. The result from this experiment showed that the output variance increases as input variance increases, and that it reaches an asymptote at large σ_I^2. The firing rate r increases as the input I increases and the CV decreases as r increases.

Abbott's second experiment studied real neurons in an artificial environment: laboratory-generated signals were used as the input to actual neurons *in vivo* (see Figure 10). Both excitatory and inhibitory inputs (g_E and g_I), at different voltages, combine to create the input I that is fed into the neuron (blue triangle in Figure 10). Through this experiment it was shown that the mean of the input affects the rate but that the variance of the input is not correlated with the variance of the output. Instead, the input variance acts more like a volume control for the output, affecting the gain of the response.



Dayan and Abbott (2001) contains more detail on this subject.

The workshop's last foray in neuroscience was through the work of Emery Brown of the Harvard Medical School, whose goal was to answer two questions:

- Do ensembles of neurons in the rat hippocampus maintain a dynamic representation of the animal's location in space?
- How can we characterize the dynamics of the spatial receptive fields of neurons in the rat hippocampus?

The hippocampus is the area in the brain that is responsible for short-term memory, so it is reasonable to assume that it would be active when the rat is in a foraging and exploring mode. For a given location in the rat's brain, Brown postulated that the probability function describing the number of neural spikes would follow an inhomogeneous Poisson process,

$$\text{Prob(k Spikes)} = e^{-\lambda(t)}\lambda(t)^{k}/k!\ ,$$

where $\lambda(t)$ is a function of the spike train and locationover the time interval (0,t). (Brown has later generalized this to an inhomogeneous Gamma distribution.) Given this probability density of the number of spikes at a given location, we next assume that the locations $x(t)$ vary according to a Gaussian spatial intensity function given by

where μ is the center, W is the variance matrix, and $\exp\{\alpha\}$ is a scaling constant.

This model was fit to data, and an experiment was run to see how it performed. In the experiment, a rat that had been trained to forage for chocolate pellets scattered randomly in a small area was allowed to do so while data on spike and location were recorded. The model was then used to predict the location of brain activity and validated against the actual location. The agreement was reasonable, with the Poisson prediction interval covering the actual rate of activation 37% of the time and the inhomogeneous Gamma covering it 62% of the time. Brown concluded that the receptive fields of the hippocampus display dynamic behavior even when doing well-learned tasks in a familiar environment and that the model, using recursive state-space estimation and filtering, can be used to analyze the dynamic properties of this neural system. More information about Brown's work may be found at http://neurostat.mgh.harvard.edu/brown/emeryhomepage.htm.

MODELING WITH COMPARTMENTS

Turning to other modeling domains, Lauffenberger of MIT proposed to the workshop participants a simple taxonomy of modeling according to what discipline and goal is uppermost in the researcher's mind:

- Computer simulation—used primarily to mimic behavior so as to allow the manipulation of a system that is suggestive of real biomedical processes;

- Mathematical metaphor—used to suggest conceptual principles by approximating biomedical processes with mathematical entities that are amenable to analysis, computation, and extrapolation; and
- Engineering design—used to emulate reality to a degree that provides real understanding that might guide bioengineering design.

Byron Goldstein of Los Alamos National Laboratory presented work that he thought fell under the first and third of these classifications. He described mathematical models used for studying immunoreceptor signaling that is initiated by different receptors in general organisms. He argued that general models could be effectively used to address detailed features in specific organisms.

Many important receptors—including growth factor, cytokine (which promote cell division), immune response, and killer cell inhibitory receptors—initiate signaling through a series of four biological steps, each having a unique biological function. Building on work of McKeithan (1995) that proposed a generic model of cell signaling, Goldstein developed a mathematical model for T-cell receptor (TCR) internalization in the immunological synapse. Goldstein's model takes different contact areas into account and was used to predict TCR internalization at 1 hour for the experiments in Grakoui et al. (1999).

To date, the major effort in cell signaling has been to identify the molecules (e.g., ligands, receptors, enzymes, adapter proteins) that participate in various signaling pathways and, for each molecule in the pathway, determine which other molecules it interacts with. As the number of participating molecules has grown and new regulation mechanisms have been discovered, it has become clear that a major problem will be how to incorporate this information into a useful predictive model.

To have any hope of success, such a model must constantly be tested against experiments. What makes this possible is the ability of molecular biologists to create experimental systems containing only small numbers of signaling molecules. Thus, separate parts of the model can be tested directly.

Where are we at the moment in our attempt to build a detailed model of cell signaling? Goldstein has used two complementary approaches, deterministic and stochastic, to create detailed models of cell signaling:

(i) An algorithm has been created to generate the chemical rate equations that describe the dynamics of the average concentrations of chemical species involved in a generic signaling cascade;
(ii) A stochastic model for the time dependence of the state concentrations has been developed, and it has been shown that the stochastic and deterministic formulations agree in the cases studied to date; and
(iii) A model has been created for the signaling cascade that is mediated by the immunoreceptor that plays a central role in allergic reactions. This model

includes a bivalent ligand, a monovalent receptor, and the first two enzymes in the cascades, Lyn and Syk.

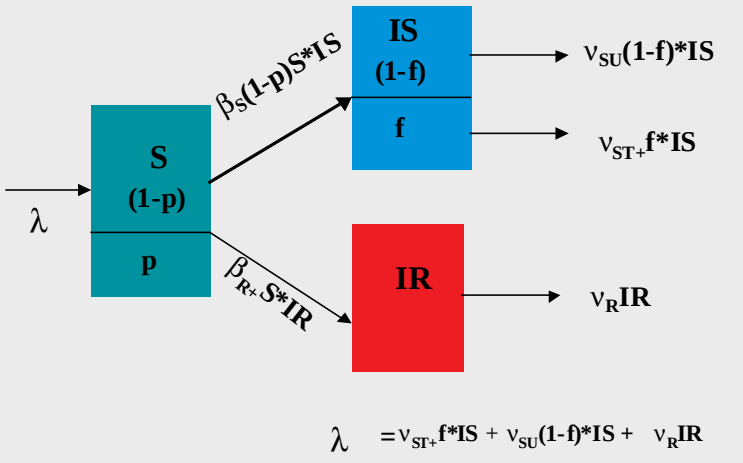
Additional information on Goldstein's modeling may be found at http://www.t10.lanl.gov/profiles/goldstein.html.

Moving from intracellular processes, Bruce Levin of Emory University presented some research that uses mathematical models to understand trends in antibiotic resistance, a major public health concern worldwide. Levin is addressing the need to know what trends are and are not of serious importance. As an example, he noted that resistance to vancomycin (an antibiotic) increased from approximately 1% in 1989 to 16% in 1997. It does not necessarily follow, though, that this is a serious problem. As is stated in Lipsitch *et al.* (2000) :

> *Although it generally is assumed that use of a particular antibiotic will be positively related to the level of resistance to that drug… it is difficult to judge whether an intervention has been successful… Mathematical models can provide such quantitative predictions, which naturally give rise to criteria for evaluating the interventions . . .*

Population dynamics can be examined with a compartment model as shown in Figure 10. The compartments represent the disease state of the individual (*S*=susceptible, *IS*= immune/susceptible, *IR*=immune/resistant). The proportion, *p*, are those under treatment, and the parameters represent the rate of movement from one compartment to another. Based on such a model, one can calculate parameters such as basic reproductive numbers and then establish rates and conditions under which the percent of resistance will increase when a proposed treatment is applied. What is often observed in public health is that the rate of resistance changes as the efficacy of the treatment changes, with high efficacy corresponding to high resistance, and the rate of resistance increases more rapidly than it decreases.

Figure 11: Compartment Model for the Epidemiology of the Antibiotic Therapy of Prophylaxis



To further investigate how a host controls infection, Levin examined *E.coli* infection in mice, where the following threshold effect has been observed experimentally: while high doses of *E. coli* kill mice, lower doses can be brought under control. A differential equations model was developed that includes this

threshold effect, and it was found to fit the data quite well. Levin's results again illustrate one of the common themes of the workshop, that a mathematical model—built on a functional premise, even if simple, and verified with data—gives us a way of quantifying biophysical processes in a way that can lead to valuable insight about the underlying structure of the processes.