Background

The brain can be viewed as a complicated computational machine. It consists of several billion neurons, each of which functions as a nonlinear, dynamical analog-to-digital filter. These neurons are highly interconnected, and they form distinct functional circuits and subsystems that are organized both serially and in parallel. Furthermore, these complex networks are modulated by feedback circuitry and by secondary non-neuronal systems. Although information is commonly assumed to be transmitted between neurons in terms of a short-term spike rate code, neural computation in the brain is widely distributed over space and time, and the principles of information coding are poorly understood.

Since Shannon's early work in the 1950's the brain has commonly been viewed as an information processing device whose function should in principle be describable in term of information theory (see Borst & Theunissen, 1999). It is clear, however, that the brain is functionally quite different from the communication systems that motivated Shannon's work. Thus far direct application of classical information theory to neuroscience data has not proved to be particularly useful. This is likely because, until quite recently, the quantity and quality of data that could be acquired in neuroscience experiments has been limited. Neurophysiological recordings could be made from only a handful of neurons, and noninvasive approaches such as EEG produced poor data that were difficult to analyze and interepret. Recent technical advances on many fronts have begun to change this picture. Neural data are now much easier to obtain due to advances such as calcium imaging and large-scale parallel neurophysiological recording. And the rapid development of functional MRI has provided a sensitive noninvasive method for obtaining data from the human brain.

These vast new sources of neuroscience data provide many opportunities for application of information theory. However, understanding the brain as an information processing device will require fundamental advances in information theory in order to account for the complex topology, extensive feedback and unknown principles of communication and coding in the brain.

Possible Topics

Potential research topics in this area include the following:

- The use of information theory to characterize the coding properties of neurons (e.g. Sharpee et al., 2004; Wu et al., 2006). A major focus of neuroscience research is to characterize the nonlinear filtering properties of neurons. The most well developed quantitative approach to this problem is system identification. However, statistical approaches based on information theory might provide much more efficient avenue to solve this difficult problem.

- Muti-channel information estimates of the capacity of multiple neuron ensembles (e.g. Aghagolzadeh et al., 2010; Bettencourt et al., 2008). Neural systems are mass-action devices whose output is influenced by the combined activity of many local units. To understand such systems requires development of new approaches for estimating the information capacity of multiple correlated, noisy channels.

- Assessment of how information is coded in spike trains (e.g. Butts et al., 2007). In vertebrates most neurons communicate with one another by discrete electro-chemical signals called action potentials. In essence this is a pulse frequency modulation code whose precise coding principles are unknown. Information theoretic analysis could provide a powerful method for assessing possible coding principles.

- The influence of neural feedback on the information capacity of the brain. Feedback is ubiquitous in neural systems. In mammalian brains the number of feedback connections to most cortical neurons is actually higher than the number of feedforward connections. The influence of feedback on information capacity is largely unexplored.

- Assessment of the “efficient coding hypothesis” (e.g. Chandler & Field 2007; Holmstrom et al., 2010; Overath et al., 2007). Evolutionary principles suggest that neural systems should be optimally adapted to the statistics of the environment, but whether this is true (and for which systems in the brain) is not known. Information theory provides a natural way to relate the statistical properties of the environment to the coding principles of the brain.

- Rate distortion theory analysis of neural transmission (e.g. Dimitrov et al., 2010). Neurons are inherently noisy and lossy information channels. The natural way to characterize such systems is rate distortion theory, but thus far this approach has been rarely applied to neural systems.

- Optimal decoding of information from the brain (e.g. Oizumi et al., 2010). One of the most exciting areas of neuroscience today focuses on decoding. Efficient decoding is critical for construction of brain-computer and brain-machine interfaces.

- Statistical biases and limitations on the application of information theory to neuroscience data (e.g. Panzeri et al., 2008). Information theory is usually applied to man-made systems where arbitrarily large data sets can be obtained. In contrast, neuroscience experiments are usually data-limited. Current information theoretic metrics much be adjusted currently to account for these limited data sets.

- Application of information theory to functional MRI data. Functional MRI (fMRI) is a rapidly developing field that has revolutionized our understanding of the human brain over the last 15 years. Because fMRI can record signals from thousands of individual locations across the brain almost simultaneously, it provides a unique opportunity for information-based analyses aimed at elucidating the function of large-scale networks.

<u>Support</u>

The Gallant laboratory at UC Berkeley will offer support for research in this area, in the form of access to raw data (neurophysiological recordings, functional MRI data sets, and guidance on access to appropriate public data repositories), general guidance on data analysis and pointers to appropriate background material.

<u>References</u>

Aghagolzadeh M, Eldawlatly S, & Oweiss K (2010). Synergistic coding by cortical neural ensembles. IEEE Transactions on Information Theory, 56, 875-889.

Bettencourt LM, Gintautas V, Ham MI (2008). Identification of functional information subgraphs in complex networks. Physical Review Letters, 100, 238701.

Borst A, & Theuniseen FE (1999). Information theory and neural coding. Nature Neuroscience, 2, 947-957.

Butts DA, Weng C, Jin J, Yeh C-I, Lesica NA, Alonso J-M, & Stanley GB (2007). Temporal precision in the neural code and the timescales of natural vision. Nature, 449, 92-96.

Chandler DM, & Field DJ (2007). Estimates of the information content and dimensionality of natural scenes from proximity distributions. Journal of the Optical Society of America A, 24, 922-941.

Holmstrom LA, Eeuwes LBM, Roberts PD, Portfors CV (2010). Efficient encoding of vocalizations in the auditory midbrain. Journal of Neurocience, 30, 802-819.

Overath T, Cusack R, Kumar S, von Kriegstein K, Warren JD, Grube M, Carlyon RP, & Griffiths TD (2007). An information theoretic characterization of auditory encoding. PLOS Biology, 5, 2723-2732.

Panzeri S, Senatore R, Montemurro MA, & Petersen RS (2008). Correcting for the sampling bias problem in spike train information measures. Journal of Neurophysiology, 98, 1064-1072.

Sharpee T, Rust NC, & Bialek W (2004). Analyzing neural responses to natural signals: maximally informative dimensions. Neural Computation, 16, 223-250.

Wu, M.C.-K., David, S. V., & Gallant, J.L. (2006). Complete functional characterization of sensory neurons by system identification. Annual Review of Neuroscience, 29, 477-505.