Principles of neuronal coding

Grand Challenges Workshop
Institute of Medicine Neuroscience Forum
25 June 2008

William Bialek
Joseph Henry Laboratories of Physics
Lewis-Sigler Institute for Integrative Genomics
Princeton Center for Theoretical Science
Princeton University

http://www.princeton.edu/~wbialek/wbialek.html
The enormous variety of our sensory experiences, abstract thoughts, and actions all are represented in our brain by sequences of identical electrical pulses, called action potentials or "spikes."

This is the problem of neural coding.
What are the elementary symbols in the code, and what do they represent?

Classical ideas about “rate,” evolution of ideas about “timing.”
Classical ideas of feature selectivity, modern version as “dimensionality reduction.”
How are features extracted (invariantly, robustly)? Coding meets computation.
*Are populations of cells more than the sum of their parts?

Why does the code have the structure that it does, as opposed to any other possible structure?
*Has the brain chosen codes that are, in some precise mathematical sense, optimal? Could this be a principle from which aspects of neural function can be predicted?

Are there principles that could unify our understanding of information flow in biological systems?

There is no way to give an exhaustive account of these efforts here, so I’ll choose some examples* (shamelessly drawn from work done by my colleagues and myself), and then point to lessons for the future.

Note: I am going to hide almost all of the math, but I give references.
Are populations of cells more than the sum of their parts?

states of the network = binary patterns of spiking and silence

<table>
<thead>
<tr>
<th></th>
<th>cell #1</th>
<th>cell #2</th>
<th>cell #3</th>
<th>cell #4</th>
<th>cell #5</th>
<th>cell #6</th>
<th>cell #7</th>
<th>cell #8</th>
<th>cell #9</th>
<th>cell #10</th>
<th>states or patterns</th>
</tr>
</thead>
<tbody>
<tr>
<td>a moment ago</td>
<td>no spike</td>
<td>spike</td>
<td>no spike</td>
<td>no spike</td>
<td>no spike</td>
<td>spike</td>
<td>no spike</td>
<td>spike</td>
<td>no spike</td>
<td>spike</td>
<td>100010101</td>
</tr>
<tr>
<td>right now</td>
<td>spike</td>
<td>no spike</td>
<td>no spike</td>
<td>spike</td>
<td>no spike</td>
<td>no spike</td>
<td>no spike</td>
<td>no spike</td>
<td>spike</td>
<td>no spike</td>
<td>1001000010</td>
</tr>
<tr>
<td>the next moment</td>
<td>no spike</td>
<td>no spike</td>
<td>spike</td>
<td>no spike</td>
<td>spike</td>
<td>no spike</td>
<td>spike</td>
<td>no spike</td>
<td>no spike</td>
<td>no spike</td>
<td>10101000</td>
</tr>
</tbody>
</table>

A common observation in many brain areas (including cortex) is that when we look at pairs of cells, the 1s and 0s are only weakly correlated.

Does this mean we can ignore correlations? Or could something dramatic be hiding in these rather mundane observations?

Analogous problems in statistical physics lead us to be suspicious of widespread correlations ...

(e.g., pairs of ganglion cells in the vertebrate retina as it responds to natural movies)
Although pairwise correlations are weak, there are collective effects in larger populations...

What can pair correlations tell us about the distribution of states in the whole network? In general, nothing. But there is a unique “simplest” or “least structured” model consistent with the measured correlations (technically, the maximum entropy model).

These minimal models are exactly the Ising model in statistical mechanics, and also the Hopfield model of neural networks. Here these “physics-style” models emerge directly from the data!

Multiple attractors, phase transitions: more than the sum of its parts!
Optimal information transmission?

In an efficient neural code, the input/output relation of neurons is matched to the distribution of inputs. Intuitively, we want the cell to be driven through its full dynamic range. This can be formalized by asking for a code that maximizes information transmission.


Unifying principles?

Mother puts mRNA for Bicoid at (future) head of the embryo

Bicoid activates expression of Hunchback

Here we measure the input/output relations and noise ... Can we predict the distributions? Same principle as for neurons!

(intensity of bcd stain) ~ (future) bcd concentration

(intensity of hb stain) (no free parameters!)

Probing the limits to positional information.
T Gregor, DW Tank, EF Wieschaus & W Bialek, Cell 130, 153-164 (2007)

Information flow and optimization in transcriptional regulation.
Theory (not just models) as a full partner with experiment

Maximum entropy for pairwise + recording many cells simultaneously leads to reconcile weak correlations with strong collective behavior; connect experiments directly to large body of neural network models; strength of correlations seems “just enough” for critical behavior.

Information maximization + measuring distributions (not means) leads to new forms of adaptation in the neural code; new questions about mechanisms; common principles across neural and genetic signaling.

Crucially, theory doesn’t just come after exp’t, it can and should come before: pointing to the next interesting experiment, and motivating larger scale, more quantitative measurements.

A grand challenge: principled (not parameterized!) theories for real networks.