Gatsby Computational Neuroscience Unit
Theoretical Neuroscience

Final Examination
29 Jan 2015

Part I – short questions

There are five sections with four questions each. Please answer three out of each four, starting the answers for each new section on a new page. Don't forget to write your name at the top of each block of answers.

You have a maximum of 6 hours for this exam.

Good luck!

1 Biophysics

1. An HCN neuron activates when hyperpolarized and inactivates when depolarized. For this neuron, when the membrane potential is subthreshold, the dynamics looks something like

$$\tau_m \frac{dV}{dt} = -(V - \mathcal{E}_L) - g_0 x (V - \mathcal{E}_{Na})$$
$$\tau_x \frac{dx}{dt} = -(x - x_\infty(V))$$
$$x_\infty(V) = \frac{1}{1 + \exp((V - V_0)/\Delta V)} .$$

   Use standard values for $\mathcal{E}_L$ and $\mathcal{E}_{Na}$, and assume that $V_0 = \mathcal{E}_L$ and $\Delta V = 15$ mV. Sketch the response of the neuron to small, brief current injections in two limits: $\tau_x \gg \tau_m$ and $\tau_x \ll \tau_m$. Consider both positive and negative current injections, and assume that the system was in steady state before the current injection.

2. Consider a quadratic integrate and fire neuron whose synaptic drive has infinitely long memory,

$$\tau_m \frac{dV}{dt} = \frac{(V - \mathcal{E}_L)(V - \mathcal{E}_L - \Delta V)}{\Delta V} - g_0 x (V - \mathcal{E}_{Na})$$
$$\frac{dx}{dt} = \sum_i \delta(t - t_i)$$

   where the t_i are the times of the presynaptic spikes. Assuming that initially $x = 0$, how many pre-synaptic spikes would it take before the neuron starts firing repetitively? Express your answer in terms of $\mathcal{E}_L$, ΔV, $\mathcal{E}_{Na}$ and g_0.

3. You are designing a brain, and you want to fill it with lots of neurons and lots of connections. As a first approximation, you can make your neurons either spherical or cylindrical. The question is: which takes less volume? (Ignore issues of wire length, dendritic nonlinearities, and electrotonic length.)

   Assume each neuron receives K synapses, and each synapse takes 5 square microns. The cylinder (which for simplicity we'll take to be of constant radius) corresponds to dendrites, so you will have to assume a reasonable minimum diameter. You will also need to assume a reasonable value for K. Be sure to tell us what they are.

4. The STDP rule can we written

$$\frac{dW_{ij}}{dt} = \eta \sum_{ij} \left[K_+(t_i - t_j)\delta(t - t_i) - K_-(t_j - t_i)\delta(t - t_j) \right]$$

   where W_{ij} is the connection strength from pre-synaptic neuron j to post-synaptic neuron i; the t_j are the times of the pre-synaptic spikes and the t_i are the times of the post-synaptic spikes; $\delta(\cdot)$ is the standard Dirac delta-function, and K_+ and K_- are temporal kernels – for instance,

$$K_+(\tau) = f_+ \Theta(\tau) e^{-\tau/\tau_+}$$
$$K_-(\tau) = f_- \Theta(\tau) e^{-\tau/\tau_-}$$

   where Θ is the Heaviside step function. Explain, mainly in plain English, what this rule does, and why it is equivalent to the usual STDP rule.

2 Networks

1. Consider Wilson-Cowan equations of the form

$$\tau \frac{d\nu_E}{dt} = A_E \phi\big(K^{1/2}(J_{EE}\nu_E - J_{EI}\nu_I + I_E)\big) - \nu_E$$
$$\tau \frac{d\nu_I}{dt} = A_I \phi\big(K^{1/2}(J_{IE}\nu_E - J_{II}\nu_I + I_I)\big) - \nu_I$$

where ϕ is a sigmoidal function of its argument,

$$\phi(x) = \frac{1}{1+e^{-x}}$$

and K is the number of connections, assumed to be large.

To determine stability, we linearize around a fixed point, denoted (ν_{E0}, ν_{I0}). Show that the linear analysis is valid (meaning quadratic and higher order terms are small) so long as $|\nu_E - \nu_{E0}|$ and $|\nu_I - \nu_{I0}|$ are small compared to $K^{-1/2}$. Show that when $|\nu_E - \nu_{E0}|$ and $|\nu_I - \nu_{I0}|$ are $\mathcal{O}(1)$, the trajectories depend only on the A's and τ (so long as the firing rates are farther than $\mathcal{O}(K^{-1/2})$ from a nullcline).

2. Consider Wilson-Cowan equations of the form

$$\tau \frac{d\nu_E}{dt} = \nu_E - g\nu_I + \epsilon f(\nu_E, \nu_I)$$
$$\tau \frac{d\nu_I}{dt} = 2(\nu_E - g\nu_I) + \epsilon h(\nu_E, \nu_I)\,.$$

Show that if $g > 1/2$ and $\epsilon \ll 1$, the system is well described by the set of equations

$$\tau \frac{dz}{dt} = \epsilon \frac{2f(gz, z) - h(gz, z)}{2g-1}\,.$$

3. A fully connected Hopfield network evolves according to

$$s_i(t+1) = \Theta\left[\frac{1}{Nf}\sum_{j\neq i}\sum_{\mu=1}^{p}(\xi_i^\mu - f_0)\xi_j^\mu s_j(t)\right]$$

where Θ is the Heaviside step function, the sum over j runs from 1 to the number of neurons, N, and the ξ_i^μ are random binary vectors,

$$\xi_i^\mu = \begin{cases} 1 & \text{probability } f \\ 0 & \text{probability } 1-f\,. \end{cases}$$

As usual, assume that N and p are large.

Let $s_j(0) = \xi_j^1$. Show that

$$s_i(1) = \Theta\left[\xi_i^1 - \mu + \sigma\eta_i\right]$$

where η_i is zero mean, unit variance, Gaussian noise, and

$$\mu = f_0 - pf(f - f_0)$$
$$\sigma^2 = \frac{p}{N}[f(1-f) + (1-f^2)(f-f_0)^2]\,.$$

Show that the expected fraction of bit flips (the probability that $s_i(1)$ is different from ξ_i^1) is

$$\langle\text{fraction of bit flips}\rangle = fH\left(\frac{\mu-1}{\sigma}\right) + (1-f)H\left(\frac{-\mu}{\sigma}\right)$$

where H is the cumulative normal function. What does this tell us about the optimal setting of f_0 in the large p limit?

4. Consider a network of point neurons described by the Hodgkin-Huxley equations. We're going to focus on two terms: the fast sodium current and the excitatory drive. We therefore write the equation for the i^{th} neuron as

$$\tau \frac{dV_i}{dt} = ... - \overline{g}_{Na} m^3 h (V_i - \mathcal{E}_{Na}) - \sum_j W_{ij} g_j(t)(V_i - \mathcal{E}_E)\,.$$

The "..." takes care of all the missing terms. There are K nonzero terms in the sum over j. The question is: how big does K have to be to overwhelm the fast sodium current? Assume an average EPSP has amplitude $\overline{V}_{EPSP}$ mV, and approximate the conductance changes with delta-functions,

$$g_j(t) = \sum_k \delta(t - t_j^k)$$

where t_j^k is the time of the k^{th} spike on neuron j. Show that

$$\left\langle \sum_j W_{ij} g_j(t)(V_i - \mathcal{E}_E) \right\rangle \approx K \overline{V}_{EPSP} \overline{\nu} \tau_m$$

where $\overline{\nu}$ is the average presynaptic firing rate. Use this to estimate the value of K at which the sodium and synaptic conductances are about equal. Assume that $m = h = 1/2$, and that V_i can be set to its resting membrane potential. Recall that $\overline{g}_{Na}$ is about 400. Your answer will, of course, depend on the average firing rate.

3 Coding

1. One model of scalar timing suggests that noisy pulses from a stochastic generator are added up in an accumulator to a threshold θ. If the generator is an homogeneous Poisson process with rate λ, what is the probability distribution of the time it takes to hit the threshold? What is the coefficient of variation (CV)? How would the CV change if the rate was drawn, independently on each trial, from an exponential distribution?

2. Define the receiver operator characteristic (ROC) curve associated with discriminating a signal source from a noise source. What is the significance of the area under the curve? Sketch its form for the case when signal and noise are both exponentially-distributed random variables with different rates. What is a formula for the area?

3. A multidimensional photoreceptor signal $\mathbf{x}$ with Gaussian statistics (mean 0, covariance matrix Σ) is subject to a noisy linear transformation by the retina $\mathbf{y} = R \cdot \mathbf{x} + \boldsymbol{\epsilon}$ where ϵ is uncorrelated Gaussian noise. What form should R take to maximize the mutual information between $\mathbf{x}$ and $\mathbf{y}$ subject to a power constraint on $\mathbf{y}$ (that $\langle \mathbf{y} \cdot \mathbf{y} \rangle < \lambda$). Will R be unique?

4. Consider a population code on a circular domain – with the activity r_θ of neuron θ in response to a presented angle ϕ being Poisson, with mean $e^{a\cos(\phi-\theta)}$. Assume the tuning curves are homogeneous and dense, so that we can replace sums over θ with integrals with density $\rho(\theta) \propto 1$, and that $\int_0^{2\pi} d\theta \; e^{a\cos(\phi-\theta)}$ can be approximated as being independent of ϕ What is the Fisher information associated with ϕ?. Note that $\int_0^{2\pi} \; d\omega \cos(\omega) e^{a\cos(\omega)} = 2\pi\mathcal{B}_1(a)$ where $\mathcal{B}_1(\cdot)$ is a modified Bessel function of the first kind.

4 Systems

1. What power do arguments based on information theoretic redundancy have for explaining cortical receptive fields (by reference to retinal ganglion cell receptive fields)?
2. What concerns might EU reviewers have if presented with a plan to simulate faithfully every neuron in the brain and thus derive a bottom-up understanding of neural processing?
3. Estimate, with reasoning, how much information from the booming buzzing confusion of the sensory world can get through to one's attention per second.
4. Vanilla learning rules – the Hebb rule, or even STDP – tend to be unstable. Explain why, and provide two ways of fixing the problem.

5 Learning

1. Consider a *supervised* Hebbian learning rule for a linear perceptron, with inputs $\mathbf{u}^m \in \{-1, 1\}^{N_u}$ and output $v^m \in \{-1, 1\}$, for $m = 1 \ldots M$. Here, each input dimension (u_i^m) and output (v^m) is set to ± 1 independently at random with probability 0.5. If the neuron's threshold for classifying the output as -1 or 1 is 0 (and N_u is big), write down an expression for the probability that it makes an error on an input $\mathbf{u}^n$, say, in terms of a cumulative Gaussian function.

2. Write one learning rule each for the direct and indirect actors in a reinforcement learning context. What happens to both your actors when one of the actions is permanently better than the others?

3. Prove that exponential discounters make temporally-consistent decisions, but provide an example proving that a hyperbolic discounter can make temporally inconsistent decisions. What sort of mechanism(s) could a latter use to mitigate the resulting difficulties?

4. Sketch on-centre and off-centre receptive fields in the retina and their rough variance-covariance functions as a function of spatial separation and spatial frequency (in 1d, if you prefer). What would happen if they were connected to a population of subtractively- or divisively-normalized Hebbian learning neurons.