Gatsby Computational Neuroscience Unit
Combined Qualifying Examination

29 Jan 2015

There are 10 questions. Answer all of them, to the best of your ability. Each is worth 10 marks. No reference materials are allowed.

It should take no more than 6 hours

Good luck!

1. Define the receiver operator characteristic (ROC) curve associated with discriminating a signal source from a noise source. What is the relationship between the area under the ROC curve and the probability of error in a two-alternative forced choice discrimination? Prove this result. Sketch the form of the ROC curve for two cases: one in which signal and noise are both Gaussians with different means and variances; the second when signal and noise are both exponentially-distributed random variables with different rates. Provide formulæ for the area under the curve in both cases (you may quote the cumulative normal probability function).

2. 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? Consider the case that $\mathbf{x}$ is itself corrupted by Gaussian white noise, so $\mathbf{y} = R \cdot (\mathbf{x} + \boldsymbol{\eta}) + \boldsymbol{\epsilon}$. Suggest intuitively what should happen to R as a function of the variance of the noise $\boldsymbol{\eta}$. Write down the calculation to show that this is indeed the case.

3. 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 $e^{a\cos(\phi-\theta)}$. Assuming the tuning curves are homogenous 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. Consider discriminating ϕ from $\phi + \delta\phi$, where $\delta\phi$ is small. What is the score function in this case? How does it relate to optimal discrimination?

4. Consider the case for reinforcement learning of deciding between two actions, A and B, associated with reward densities $p_A(r)$ and $p_B(r)$ respectively. Define direct and indirect actors and formalize what each is intending to achieve. For the direct actor, show the relationship between this intent and the learning rule. Consider the case for observational learning in which samples of rewards for both A and B are provided whichever was chosen. What is the formula for how the variances of the components of the indirect actor evolve over learning trials?

5. The inspection game used by Dorris & Glimcher (2004) to investigate choice and the parietal cortex involves me (a worker: I can work (w) or shirk (s)O and you (a manager: you can inspect (i) or laze (l)). The payoffs to both of us are:

payoff to:	wi	wl	si	sl
me	50	50	0	100
you	200-I	200	100-I	0

where I determines the cost of inspection. Define a (mixed strategy) Nash equilibrium in this context, and find its form as a function of I. When might it be hard for regular, non-game-theoretic, indirect reinforcement learning actors to learn to play this game, and why?

6. 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. Assuming that these are spread uniformly over the input (and so act as convolution filters), what are the eigenvectors of these covariance - in both 1d and 2d cases (note the two structurally different classes in 2d). If the outputs of these filters are input to a layer of subtractively or multiplicatively-normalized Hebbian learning, what would happen, and why?

7. Consider a quadratic integrate and fire neuron driven by a Poisson spikes,

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

where the t_i are the times of the presynaptic spikes. Assume the spikes are Poisson with rate ν. Estimate the value of ν that will make the neuron fire. You may assume that g_0 is small. Be clear about what approximations you make.

8. You are designing a brain, and you want to fill it with lots of neurons and lots of connections. Discuss the advantages, with respect to volume, of having dendrites versus spherical neurons. Estimate how large your brain would be if all neurons were spherical, but it retained its connectivity. Ignore issues of wire length, dendritic nonlinearities, and electrotonic length.

9. 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. All the J's are positive.

Draw the nullclines. Assume

$$J_{EI}J_{IE} > J_{EE}J_{II}$$
$$0 < \frac{I_I}{J_{II}} < \frac{I_E}{J_{EI}}$$
$$A_E > A_I\,.$$

Sketch the nullclines. When the firing rates are farther than $\mathcal{O}(K^{-1/2})$ from the nullclines, write explicit expressions for the firing rates as a function of time (those expressions will depend on the A's, τ, and initial conditions). Sketch representative trajectories.

10. 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

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

Assume, as usual, that N and p are large. Let that $s_j(0) = \xi_j^1$. Compute the expected fraction of bit flips (the probability that $s_i(1)$ is different from ξ_i^1). Plot the expected number versus f_0.