TNI, 2010

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 150 minutes for this exam. Please mark any answers provided after 60 minutes.

Good luck!

1 Biophysics

1. What happens to time constant of a passive neuron if it changes from a sphere to a cube without changing its volume? Explain your reasoning.
2. Explain how NMDA channels act as coincidence detectors of pre and postsynaptic spikes.
3. What are the boundary conditions on voltage and axial current at a branch point in a dendritic tree?
4. A passive neuron is gap-junction coupled to an active one, and so obeys the equation

$$CdV/dt = -g(V - \mathcal{E}_L) - g_1(V - U)$$

   where U is the membrane potential of the active neuron. Assume that the active neuron emits a spike at time $t = 0$, which we approximate as $U(t) = \mathcal{E}_L + \Delta V \tau \delta(t)$. What is $V(t)$ in terms of all the parameters of the model?

2 Networks

1. Consider a Hopfield network with the update rule

$$s_i(t+1) = \text{sign}\left[\frac{1}{N}\sum_{j=1}^{N}\left(\sum_{\mu=1}^{p}\xi_i^\mu\xi_j^\mu\right)s_j(t)\right]$$

   where $\xi_i^\mu = \pm 1$. If the ξ's are chosen so that $\sum_{i=1}^{N}\xi_i^\mu\xi_i^\nu = N\delta_{\mu\nu}$, show that the network can exhibit perfect recall up to $p = N$.

2. Consider a discrete time update rule of the form $\nu_i(t+1) = \phi(h_i(t))$ where the ν_i are firing ates, ϕ is a sigmoidal gain function (which you won't need) and

$$h_i(t) = \sum_{j=1}^{N} w_{ij}(t)\nu_j(t)\,.$$

   Let $w_{ij}(t) = w_0/N + \delta w_{ij}(t)$ where w_0 is constant and $\delta w_{ij}(t)$ is an *iid* random variable that is pulled, on each time step, from a distribution with mean zero and variance σ^2/N. Assuming that $w_{ij}(t)$ and $\nu_j(t)$ are uncorrelated, what is the mean and variance of $h_i(t)$?

3. Consider a randomly connected network of excitatory and inhibitory neurons operating in the balanced regime. What would happen to the average excitatory and inhibitory firing rates if you increased slightly the external excitatory drive to all the inhibitory neurons? Explain your reasoning using either nullclines or equations (I would suggest the former).

4. You are reviewing a paper in which the authors simulate a randomly connected network of excitatory and inhibitory neurons. The network contains 100 neurons, the connection probability is 10%, the PSP size is about 1 mV, and there are no inhibitory-inhibitory connections. Central to the paper is the claim that the network will exhibit realistic firing patterns if the number of neurons is scaled up to 10s of thousands without changing the PSP size or connection probability. Why should you reject the paper?

3 Coding

1. A particular synapse invariably fails to release a vesicle for every **second** presynaptic spike. Assume that the presynaptic spike-train is drawn from a homogeneous Poisson process. What is the interval distribution between synaptic releases? Derive the expression. What about if each spike causes a vesicle to be released with a probability of 0.5 (with release to one spike independent of the release to any other)?

2. Many methods for characterising neuronal response properties can be seen as algorithms that search for ways in which the spike-triggered stimulus ensemble differs from the ensemble of all presented stimuli. What measures of ensemble difference do each of the STA, STC and MID methods exploit?

3. Consider the classic signal detection theory setup: two stimulus values s_0 and s_1 correspond to two different one-dimensional response distributions $p(r|s_0)$ and $p(r|s_1)$ such that $E[r|s_1] > E[r|s_0]$. Consider a simple two-interval two-alternative classifier: given one sample response from each of these distributions the classifier compares their values and reports that the larger one corresponds to s_1. Show that the probability that this classifier is correct is equal to the area under the ROC curve associated with the distributions.

4. Give three arguments **against** the idea that population codes have evolved in part to maximize the Fisher information in neuronal responses to single (scalar-valued) sensory stimuli.

4 Systems

1. Reducing the ambient light in a room often helps to resolve detail in low-contrast slides. But, as photons from the projector simply add to photons from other light sources, the absolute difference in photon count from the two similar-colour regions on the screen is the same under both high- and low-ambient light conditions. So why is it easier to see with the lights down?

2. Explain (or sketch) the pattern of connections between retina, LGN and cortex that ensures that information from each half of the visual field is sent largely to the opposite cerebral hemisphere.

3. Suggest two possible roles that might be served by the thalamic relay stage in the neocortical sensory pathways for vision, audition and somatosensation.

4. Would a rational subject ever experience unexpected uncertainty? Justify your answer.

5 Learning

1. Consider the case for conventional Hebbian learning ($\Delta w_i = \alpha v u_i + \text{normalizing factors}$ that the output $v = \sum_i w_i u_i$, where $u_i > 0$ have the property that $\sum_i u_i = 1$ for all patterns, and we seek to normalize the weights so that $\sum_j w_j = 1$. Write down the version of the learning rule that: (a) normalizes this subtractively; (b) normalizes it multiplicatively, **in both cases so that the weights remain exactly** $\sum_j w_j = 1$.

2. Consider a neuron $v(\mathbf{u})$ reporting on an input vector $\mathbf{u}$ drawn from a distribution $P(\mathbf{u})$. If one seeks to minimize the average squared reconstruction error where reconstruction $\hat{\mathbf{u}} = \mathbf{g}v(\mathbf{u})$ is linear, i.e., to minimize $\langle |\mathbf{u} - \mathbf{g}v(\mathbf{u})|^2 \rangle$, is there any reason for the mapping $v(\mathbf{u})$ to be non-linear? Justify your answer.

3. Define exponential and hyperbolic discounting of future rewards. Why does the latter give rise to preference reversals?

4. Consider an agent following a Markov process with transition matrix P_{xy} and rewards r_x, and having a regular exponential discount factor γ. If the value function is defined as $V_y = \langle \sum_t r_{s_t} \gamma^t \rangle_{s_0 = y}$, and the agent starts at $s_0 = x$, what are the expected values of:

$$r_0 + \gamma V_{s_1}$$
$$r_0 + \gamma r_1 + \gamma^2 V_{s_2}$$
$$r_0 + \gamma r_1 + \gamma^2 r_2 + \gamma^3 V_{s_3}$$

   TD(λ) combines all of these to estimate V_x - why might this be a good idea?