Problem set 6 Computational Neuroscience
(Hand in Problem 1 on October 22nd)

Problem 1. Determining the neural code is one the key problems in systems neuroscience. Two open (and contentious) questions are: do spike times matter or is the firing rate estimated across a sufficiently small time interval enough; the other is whether correlations between neurons contain useful information. Information theory can help answer these questions by calculating the mutual information I(R;S) between the response R and stimulus S, and compare coding schemes where the information in the precise spike timing (or correlations) is erased by shuffling spike times or across neurons. However these procedures not only alter the actual mutual information, they may also affect the bias, thereby confounding the interpretation of experiments with inadequate sample sizes. Another problem is what is meant by correlations and what is meant by containing information. I want to you to read a couple of papers with these issues in the back of your mind and formulate an opinion on the question whether correlations play a role and/or ways to measure them. There are no definite answers to this question.

a) Nirenberg et al, Nature 2001, Volume 411, p 698
b) comment by Markus Meister
c) explanation of the definition by Nirenberg & Latham, Proc. Nat Acad. Sciences 2003, Vol 100, p 7348.

Problem 2. *Coupling between features in Self-organized maps (SOM) (or Kohonen maps after the originator).* Kohonen maps are a way to perform dimensionality reduction, which is, for instance, necessary in the primate visual cortex where an orientation map, retinotopic map, ocular dominance map, and perhaps more needs to be represented on a 2-dimensional cortex. This makes it difficult to map all features uniformly, leading to singularities such as pinwheels and non-uniform gradients in feature values. The question is whether there are correlations between gradients of different features, for instance, it could be that when the orientation varies quickly on the cortex, retinotopy varies less. These effects have not been clearly observed in experiments but are predicted by SOM. SOM are calculated as follows.

X is a vector in feature space (high dimensional), r is a vector in representation space (low dimensional) and each discrete r value has associated with it an example vector m_r. The on-line algorithm presents samples X_t which are used to adapt the m_r so that they cover feature space. The total number of r values is denoted by N.

Step 0: assign random values in feature space to m_r.
Step 1: find the index r* of the m_r closest to X_t
Step 2: shift all m_r with r values close to r* according to

$$m_r \to m_r + \alpha \exp(-d(r,r^*)^2 / 2\sigma^2)(X_t - m_r)$$

repeat 1 and 2 until you feel it has been enough.

This algorithm is simple, but key decisions are: how much do I move m_r, α, how fast does the amount of change fall off with the distance from r*: σ and how big is the neighborhood of r for which I update m_r. For $d(r,r^*)$ one usually takes the Euclidean distance. Generally α and σ are reduced in time, so that algorithm converges.

a) Implement the SOM algorithm for a 1-dimensional representation and 2-dimensional feature space comprised of a rectangle with sides a and b. r then corresponds to a linear index, say "i".
b) Use this algorithm for a=b=1 and N=10, 50, 200. Explore how fast α and σ should decay in time. Comment on the effect of increasing the number of points N in representation space.
c) Repeat for a=1, b=4. Comment on the shape of the line in **feature space** defined by m_r, when r goes from the lowest to highest value in **representation space**.
d) The difference between b) and c), is that in c) the "b" direction uses more points in representation space, which can be quantified in terms of the gradient (derivative) of m_r (please show this). In the Yu et al paper (Neuron, 2005, Vol 47, p 267) the same problem is studied but for 2-dimensional representation space, and 5 or more dimensional feature space. They address the issue that Kohonen maps predict correlations between gradients in feature values, but that these have not been found experimentally so far.

> 1) Describe the three ways in which these correlations are quantified in the paper.
> 2) Why is the Ferret cortex useful to address this issue experimentally?
> 3) What is their explanation for the fact that the correlation in gradients between orientation and ocular dominance is still rather low for the Ferret? How about the lack of correlations found in experiments in other species?
> 4) Think of issues on which you could criticize this paper.