Computational Neuroscience
Dr. Sharba Bandyopadhyay
Department of Electronics and Electrical Communication Engineering
Indian Institute of Technology Kharagpur
Week – 12
Lecture – 58

Lecture 58 : Optimal Coding of Deviant Stimuli in Development

Welcome. So, we are in our last week of computational neuroscience course and in the previous two lectures we have covered some of the theoretical principles based on sparse coding and or efficient coding as to how we can see what receptive fields are optimal in a way that there is minimal energy usage and the representation of the stimuli are also adequate in terms of responses of neurons. So, following this we will be discussing few more theoretical aspects along with modeling aspect with along the same lines again that is maximizing mutual information or information transfer and with minimal energy that is representing our stimuli with minimal energy and with those ideas in one case we will show how neural system develops in a particular way and in another case we will see how on theoretical basis we can come up with spike timing dependent plasticity kind of rules. So, let us go to our examples that we have done so far with long term plasticity and long term depression either short term plasticity both of them together we will consider one example this is from one of our own work and so if you recollect we had said that there are set of neurons called sub plate neurons in the cortex in the cortical region which are the first born neurons and they are transient in nature. Their first born neurons in the cortex and they are transient in nature in the sense that they die after a certain period of time and their existence is coincident with what we know as critical period. So, in one of our work we show that in the auditory system earlier from earlier work we what you already know is that during critical period if this is the primary auditory cortex surface in the cortical regions let us say this is A 1 we know that it is arranged tonotopically from low frequencies to high frequencies with medium frequencies in between.

So, there is a gradual gradation so that is what we call as tonotopy. We also said that if during critical period we expose an animal with presentation of one particular frequency that is represented normally let us say over this hatched region let us say this is 7 kilohertz and let us say this is a rat A 1 and it is presented with 7 kilohertz pips of tones simple pure tone continuously like every half a second or so. Then when the so this is the let us say animal during at birth then it goes through a critical period when that 7 kilohertz is presented and then in when the

animal is an adult the auditory cortex is mapped and this is the primary auditory cortex and compared to the actual the 7 kilohertz region gets expanded. So, this is the low frequency at high frequency and at the cost of representation of the nearby frequencies which are taken over by this 7 kilohertz region.

And if this 7 kilohertz presentation is done after critical period then that does not lead to such plastic changes. So, the idea that was pursued by us was that actually at an earlier period which showed that the sub plate neurons are actually responsive to sound and maybe not the layer 4 neuron. So, if you recollect the circuitry we have a thalamic input going into the sub plate neurons and its collateral goes on to the input layer, layer 4 of the cortex and sub plate then projects on to layer 4 and this after completion of the thalamocortical development ends up being a synapse on layer 4 of a certain weight. So, remember that this is a very minimal representation of the circuitry we had also mentioned to you that the actual circuitry has sub plate to layer 4 convergent inputs and similarly from thalamus to sub plate convergent inputs and thalamus to layer 4 convergent inputs as well. So, this is layer 4 this is sub plate and this is thalamus where this axis is the frequency axis or the tonotopic axis and many neurons many sub plate neurons in along the frequency axis project on to layer 4 and many thalamic neurons along the frequency axis project on to sub plate neurons as well as layer 4 neurons and a similar kind of arrangement plastic development of the thalamocortical circuitry can be seen.

There are further complications here with inhibitory neurons playing on another role plus there are feedback from layer 4 to sub plate and sub plate to thalamus. So, things are not as simple as we represent here, but the principles that we will talk about are going to be similar. So, considering this circuitry this simple circuitry which transforms into simply thalamus to layer 4 in that scenario what is found is that if we have layer 4 neurons if we record from sub plate neurons these sub plate neurons tend to be deviant selective that is they are deviant detectors whereas, layer 4 neurons are not at ages even before what is considered as critical period. So, in other words what is found is in the auditory cortex if you record from sub plate neurons at ages of P 10 that is 10 day old mice or P 9 9 day old mice this is before critical period which is believed to be P 12 to P 16 for the 7 kilohertz experiment that we talked about it is this period when the 7 kilohertz presentation works to have an over representation of 7 kilohertz. So, preceding that period in P 9 P 10 P 11 in those regions even P 8 what is found is that the sub plate neurons are deviant selective in the sense that if you play a standard frequency tone as we have discussed earlier and deviant tone at low probability it is found that the sub plate neurons are hardly responding to the standard, but they

respond to the deviant token with a larger response rate much larger response rate and there is hardly any response to the standard.

So, that is maybe initially when you present the standard there is a blip response in the very onset, but beyond that it is simply adapted to the standard sound and there are no responses whatsoever to the standard and it only responds to the deviant or at least that is what most of the data says. So, in other words if in these ages if we do the 7 kilohertz experiment that is at P 9 to 10 P 9 to P 10 or P 11 before critical period. So, let us say this is critical period before this if you do the 7 kilohertz presentation as discussed then based on these data that we are talking about standard there is no response then the sub plate neurons are not at all driven by this 7 kilohertz continuously presented tone and so there is no effect of a continuous presentation of the 7 kilohertz in terms of plasticity in the adult. So, ah it is it is then likely that the experiments that suggested that presentation of 7 kilohertz at critical period ah leads to expansion of 7 kilohertz and presentation of 7 kilohertz after P 12 to 16 does not do anything similarly the presentation of 7 kilohertz continuously before critical period also does not do anything and so that critical window ah critical period window came up to be P 12 to P 16. So, ah it is the experiment that with standard and deviant where the deviant let us say is now 7 kilohertz and the standard presentation is let us say 4 kilohertz or 17 kilohertz.

So, in that case we have an experiment where it appears that something that is being presented at a much lower probability that is actually driving the system and the standard sound is not able to drive the system. So, in this case what what is observed is that during before this critical period if we do a standard deviant presentation then later on and in as an adult we find that the responses to the deviant sound is altered that is the responses to the deviant sound is increased in the adult compared to a typically developing animal which is not exposed to any different protocol during its development. So, if the deviant is 7 kilohertz then there is a is a strengthening of the synapses that drive 7 kilohertz in the layer 4 region. The idea here is that the same thing that when if we now consider that there are 2 inputs on to the sub plate and layer 4 and consider competition ah between the 2. So, let us say this is the standard and this is the deviant frequency that is presented then the standard is unable to drive the layer 4 neuron and so this synapse weakens ah and the deviant frequency 7 kilohertz in this case the deviant this synapse strengthens.

So, as opposed to a normally developing one where there is equal presentation of S and D gradually or in fact just ah just regular development where you have some activity in both standard and deviant unlike a continuous activity in standard with minimal activity at periodically in deviants ah that ah compared to that case

scenario this synapse gets much stronger which is what is observed in the adult. So here ah there are 2 things in the modeling at play one is the modeling of the short term ah plasticity which is at the synapses that you know how to implement and there is long term plasticity also in the same synapses ah which lead to the strengthening or weakening of the standard and deviant synapses. So, these so the if you want to implement this whole idea you can actually represent actually do it in a network also and show ah this phenomena happening that the deviant representation becomes stronger and stronger in the adult with again with leaky integrator and fire sub plate neuron and layer 4 neuron and synapses that have the 3 state model implementing the short term plasticity and if we have hebbian spike timing dependent plasticity at the synapses. So this is an example of where we implement both kinds of plasticity simultaneously in the same circuit. So where is the theoretical aspect of this whole idea by the way so these are supported further by a control experiments with standard deviant presentation in the critical period and only standard presentation in the ah pre critical period and so on.

So coming to the theoretical aspect of this whole idea ah what what what is addressed what was addressed is that why is it that the deviant representation should increase. Why is it that this deviant representation should strengthen. So this problem can be formulated ah in this manner that you have a neural system and there are responses of the output let us say the layer 4 neuron that has responses to standard and responses to the deviant and in the input what we have is ah presentation of standard and presentation of deviant with different probabilities that is ah let us say this is 10 percent and this is 90 percent or more generally this is p and this is $1 - p$. So starting with very similar values of R_s and R_d in in before the plasticity takes place the idea was to ah optimize ah mutual information and ah firing rate. So in the same way from the input if this is x to the output this is y which takes on values of R_s and R_d there are 2 ah 2 kind of mean responses to ah the 2 stimuli and x is the input with the 2 stimuli s and d with probabilities as given $1 - p$ and p.

ah There is a mutual information between $I_{x,y}$ ah the input and output ah where y is the rate response and x is the stimulus identity with associated with different probabilities and the rate response also has different probabilities. And there is ah total amount of spiking activity that is $R_s + R_d$ ah in a weighted manner that is $p \times R_s$ ah $(1 - p) \times R_s$ and $p \times R_d$ this is the total rate where R_s and R_d are the mean rates in response to s and d during development. So these are variable. So and ah so of course, we need a ah need a normalization factor between the 2 we can put in a lambda and so we want minimal amount of rate and maximum amount of mutual information with ah monotonic change in R_s and R_d. So and

with a minus sign we want to ah minimize this and maximize this overall term needs to be maximized.

And with algebra finally, you can show that it indeed if we want to maximize the mutual information under certain constraints we will find that R_d has to strengthen much faster than R_s. In fact, R_s has to keep on going down to to a point where the mutual information is max is increasing whereas, the rate overall summed activity is reducing. So, this simple the the simple explanation the intuitive explanation is that well in order to maximize the mutual information between $I_{x,y}$ we need to have larger and larger separation between R_s and R_d. So obviously, if we need to maximize this with minimal amount of activity R_s has to be less than R_d it is because R_s is weighted way more times than R_d. So, it is a very very simple intuitive explanation although it requires a lot of algebra because of the framing of the problem.

So, I hope that here you appreciate the overall points that have been covered and that is let us say. So, in this case we have implementation of long term plasticity and short term plasticity both of them simultaneously we have so far been talking of them separately. Secondly, we have an example which appears counterintuitive in the sense that we always talk of learning and plasticity as something that is repeated multiple times in association with something else that association is the longer lasting one that is the one that gets potentiated. However, counterintuitively here we actually have something that is low probability that is occurring much more much much more rarely or very very less frequently that particular sounds or that particular stimulus is representation is getting strengthened. So, the low probability event is potentiated.

Thirdly, we have a theoretical way of looking at the same problem and showing that actually with mutual information maximization and minimal activity we can actually show this counterintuitive kind of plastic development in the cortex. So, with with this particular example and the theoretical way of dealing with these problems, we will be ending our lecture for this section and in the next section we will be talking about another example of deriving rules of plasticity particularly certain spike time independent plasticity rules based on completely theoretical aspects. Thank you.