Computational Neuroscience
Dr. Sharba Bandyopadhyay
Department of Electronics and Electrical Communication Engineering
Indian Institute of Technology Kharagpur
Week – 05
Lecture – 24

Lecture 24 : Stimulus to Response mapping(Coding) - I

Welcome. So, we will continue our discussion on building models of receptive fields of neurons of sensory neurons particularly. So, the problem that we are faced with is that we have a stimulus S and it is being transformed to a response R and the stimulus is could be parameterized by physical parameters some vector θ and response could be based a vector also with a certain resolution of time and as a function of time. So how is this stimulus related to the response or how is the stimulus encoded in the response is one side of the problem and similarly the other side of the problem is given a response can we predict the stimulus or can we find out what stimulus was played and it can be for other things as well not necessarily a stimulus as we have said it could be other phenomena that is going on at the time of the responses or at the time of the activity that we are observing. So this backward problem or the decoding problem is can be approached in a number of ways and the encoding problem also can be approached in a number of ways and the most simple way in which we can look at the stimulus response relationship going from S to R is by using simple linear models or linear time invariant models. So if we recollect how we have looked at the how we have developed the idea of spiking by neurons we can actually say that we have sort of box from the stimulus world let us say we have $x(t)$ as the stimulus and ultimately what we are observing is a set of spike trains or a set of spikes as a spike train.

So this we will write as $P(t)$ or we can we will essentially write this as a sum of impulses that is $\delta(t - t_i)$. So where i is varying from 1 to N_T that is what we are saying what $\delta(t - t_i)$ is the direct delta function we will go into details of it in a little bit. We are representing each spike at time t_i with this sum. So if there are there is a time window T here and we are saying that there are N_T spikes and each of those spikes are occurring at time point t_i then this is the representation of $P(t)$ that is the response.

Now in here what is going on is a very big question but for us is for us we will try to do this that how well can we describe this black box in such that given new $x(t)$ s how well can we predict the response $P(t)$ s. And so in order to build these ideas in order to get to the stage where we will be defining models within that box or understanding what is going on within that box we need to take a little bit of background even for the linear models. So the ideas that we will be using is what we call is linear time invariant systems. So as the words suggest linear we sort of give you an idea what linear is we will define it properly. Time invariant simply means that the system itself is not changing with time which is not true for neurons as we have seen that there is short term depression short term plasticity as well as long term changes and dependence of spikes due to refractory and so on.

So this is not a time invariant system however we may be able to approximate it at a particular situation as a time invariant system. And what we will see is that these kind of linear time invariant like descriptions of neurons actually works well especially when we are considering neurons in the periphery that is as we have discussed in the lateral geniculate nucleus or in the retinal ganglion cell responses or in the auditory nerve or in the cochlear nucleus certain types of neurons in the cochlear nucleus can be well described by the linear time invariant system. Although in the true sense they are not linear time invariant systems. So another reason why we want to use this kind of system approach is that we can always approximate system as a linear system over a small region of parameter space as we have done in the nonlinear phase plane analysis. We looked at the linearization around the equilibrium points.

This also similarly can be thought of linearization around a particular point in the parameter space and that description can be a linear time invariant system that will be sufficient. And as we change that baseline point in the parameter space we will have different descriptions of the system. So that is another way to conclude that well the linear time invariant system approach is actually useful even though the entire neuron is not necessarily so. Further the probably the most important idea is that even though we have all these problems with these approaches if we consider the changes in a system. So let us say we describe the neuron with a linear time invariant system approach in the initial stages.

Let us say there is some perturbation or learning or something in the system and the neuron changes its properties and then we describe the system with this linear time invariant system approach. And we see that whatever we can conclude from this approach before the perturbation or learning and after the perturbation there will be many pointers that will be available in terms of the mechanism by which the system change from the one initially which was approximated by that linear system to the new linear system. So the way the linear part of the system is changing that itself is very instructive in order to understand what underlying mechanisms may be going on that produce that change. So we will also discuss at a point of time examples of this where this approach tells us about the underlying changes in a system that is actually learning something during behavior. So with that sort of background on why we want to study linear time invariant systems to describe neurons I hope you are convinced that even though neurons are very nonlinear I mean very nonlinear is not a very good term although neurons are nonlinear still we want to use such approaches.

So now to formally say what is a linear system so first of all let us say what is a system. System is essentially what we drew in the previous slide that this is our system that takes an input and produces an output x_t as input and output is $p(t)$. So if we say that this x_t is the input we have a system s and the output is $y(t)$ for now is easier to say then the system s is called a linear system if for an input $x_1(t)$ the system produces an output $y_1(t)$ and if for an input $x_2(t)$ the system produces an output $y_2(t)$ then if the system is given the input scalar times $x_1(t)$ plus another scalar times $x_2 t$ as the input the output can be said to be $ay_1 t + by_2 t$. So this is what we call the superposition principle and if a system follows this superposition principle we say that it is a linear system. So like the example that I was giving in an earlier lecture that orientation tuning of a neuron that if we have one bar and response we know and another bars horizontal bars response we know then if we have these two together then the sum of the responses is the response when we have the two together.

That is essentially what superposition principle says and what time invariant means is that if I have an input x_t for the system s and there is an output y_t then if we shift the input by time τ then the output is also shifted by time τ and for any τ. So this is time invariance that is essentially if I give an input now and measure the output now and an hour later I give the same input I should get the same output that is our x_t translated by one hour it produces y_t also translated by one hour. So that is what we mean by time invariance and a system that follows both superposition principle and time invariance is what we call an LTI system or linear time invariant system. Now the advantage of an LTI is that we can completely describe the system if let us say the system s is now an LTI then if we have an input x_t we can always find out what y_t is if we know some particular property of the system which we call the impulse response. That is when x_t is $\delta(t)$ that is the Dirac delta function that we had said then the output y_t that we measure let us say it is h_t.

So for input $\delta(t)$ we are getting an output h_t. So if we know this h_t for an LTI we can show that for any given x_t as input to the system we can find out what the output y_t is going to be and that is $\int_{-\infty}^{\infty} x_t(\tau) h_{t-\tau} d\tau$. This is also called the convolution of x_t and h_t. So this integral is the representation of the convolution.

So why we can get that it is it follows simply from the linear time invariance property and the fact that the Dirac delta function can be used to represent any signal x_t or scaled and shifted versions of delta functions added together can be used to represent any stimulus x_t and since the system is LTI or linear time invariant and we know the output for $\delta(t)$ which is h_t we know the output for any x_t because x_t is simply scaled and shifted versions of $\delta(t)$ and linear time invariance says that we can simply sum up the responses to get the sum up the scaled and shifted sum of the responses in the same manner and get the overall output.

So to define $\delta(t)$ let us say we have been saying this Dirac delta function so this will come up later on also the $\delta(t)$ can be defined in this manner it is an impulse many of you may already know this $\delta(t)$ is that it is actually 0 for all $t \neq 0$ and it is not 0 for $t = 0$ and $\int_{-\infty}^{\infty} \delta(t)dt = 1$. There are many properties of this $\delta(t)$ that will come into use is that if we multiply $\delta(t)$ with some other x_t and integrate this over the interval let us say $-\infty$ to ∞ this simply means that it is an evaluation of x_t at wherever this delta is wherever this t is 0 so it is essentially x_0 and if we shift it if it is $t - \tau$ then it is an evaluation of x_t at τ. So these things will come up in our later lectures when we talk about how to analyze these linear time invariant systems. So having said all these things it is essential that to describe an LTI we need to be able to find out this h_t or the impulse response. So if we can somehow measure or find out what the h_t is we are done we are done describing the system.

So how do we describe or find out h_t so and how do we bring it into the overall picture of stimulus to response scenario for a neuron. So as we had been saying let us say this is the stimulus x_t and we had this big box and we have this output p_t and remember that p_t is this summation of $\delta(t - t_i)$ at for $i = 1$ to N_T that is there are N_T number of spikes in T time window for which the response for which the stimulus is on that is x_t is played. We cannot say that this itself is the LTI or we cannot approximate this as the LTI because from x_t from any x_t being able with an impulse with the impulse response it will not be possible to generate impulses in this case given a smooth kind of x_t. So now we bring into the picture what we have learnt about the neuron so far. So we said that we can represent the response by rate as a function of time or λ as a function of time and that λ_t or R_t is the driving function for an inhomogeneous Poisson process or the point process that the spike train is and these events are represented by the $\delta(t)$.

So λ_t essentially is saying that the average rate at a particular instant t that based on many many repetitions the spiking events in that instant t and δ window around it that the probability of spikes occurring in that window is proportional to this λ_t provided that there is only one spike possible in that window. So what we can approximate the system by is so $\lambda(t)$ is the driving function. So let us say we have somehow get a $\lambda(t)$ that serves as an input to a point process generator that is producing the spike train PT. So I am starting from the back end of it. We know that PT is a set of events successively occurring in time and $\lambda(t)$ is the overall driving function which is creating those events over time and that this process is an inhomogeneous Poisson process that is a λ that is changing over time.

So what is it that is actually driving the neuron to produce spikes at a particular window. It is its membrane potential when it crosses a threshold then there is a spike. The higher it crosses the threshold the longer it will take to come back down and so there will be larger period of time over which there will be high probability of spiking that is if we repeat it from over and over again with some additional noise and so we can represent this $\lambda(t)$ to be generated by the membrane potential of the neuron. So as a proxy for $\lambda(t)$ we may now think of it in this way that if we have the membrane potential somehow coming into this box that produces $\lambda(t)$ then this box would need to have a function that is converting let us say this is $y(t)$ that is converting this $y(t)$ into $\lambda(t)$ and how so the first step is that there is a threshold. So let us say our $y(t)$ is here and $\lambda(t)$ is here.

So the rate is 0 up to the threshold membrane potential. So there must be a threshold and after that there is an increase in firing rate or the probability of spiking as the membrane potential is larger and larger and we also know that there is an upper limit to this firing because of relative refractory and absolute refractory and also naturally there are other limits to the number of spikes a neuron can produce in terms of the ion channels that are present and so on and so there must be some sort of a saturation and an often used function is like this that has a threshold that saturates or a sigmoid like function or something like this that is the transformation from $y(t)$ to $\lambda(t)$. However, it is not necessary that it is always going to be like the figure that I have drawn there may be cases where this function is something like this or there may be cases where this function is simply like this straight going up to a certain range and is undefined beyond that and so on. There are many possibilities. So and that depends on maybe other inhibitory inputs that are there and so on.

So in general we can think of this $y(t)$ to $\lambda(t)$ transformation as a static nonlinearity that is we plug in y into that function and get out λ from that function and that is a nonlinear function. So let us say that we represent this by a nonlinear function that is producing λ as a function of y λ as a function of y or we can call this maybe some nonlinear function $s(t)$. Now the question is whether before $y(t)$ the transformation from $x(t)$ to the membrane potential $y(t)$ can we represent that as a linear time invariant system. So this here is preceded by a box which has which is which we are assuming to be a linear time invariant system that takes

$x(t)$ as input and convolves it with the impulse response which let us say produces something that is akin to the membrane potential of the neuron which is then passing through a static nonlinearity which is generating the probability of spikes that is the rate of spikes $\lambda(t)$ which is driving the point process. So this is the setup of the problem in the sense that if we were to approximate the transformation of the stimulus up to the neuron that we are recording from its membrane potential that entire chain of events with a linear time invariant system.

So obviously that it is not a very good I mean it will appear to be not a very good assumption but as I said we will also see that many neurons can be well described by this $h(t)$ and for the other reasons that we have mentioned this these kind of models are useful. So now we are left with the task if we want to find out what PT is going to be given any $x(t)$ we need to be able to find this nonlinear function actually and $h(t)$ actually if we know $h(t)$ that nonlinear function can be easily estimated from the available observations. So we this whole problem boils down to finding $h(t)$ since it is a linear time invariant system we know that if we know the impulse response then we know everything there is to that LTI. So if we can somehow given some $x(t)$ we know what the PTs are from these observations can somehow find out what this $h(t)$ is going to be is then we will be done with creating this model. So that we will see is will essentially be what we call the spike triggered average and that will require our ideas of autocorrelation function and cross correlation functions that we have introduced and we can show that indeed we can in spite of this nonlinearity present arbitrary nonlinearity present we can indeed model the system with an $h(t)$.

So we will take up that in our next lecture. Thank you.