

Demystifying the Brain
Prof. V. Srinivasa Chakravarthy
Department of Biotechnology
Indian Institute of Technology, Madras

Lecture – 15
Memories and Holograms Segment I

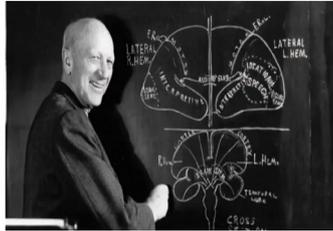
This lecture is about memories, what you call yourself is created by your memories. Your personality, character, yourself all these things are constructed out of your memories. When you say your name, you are operating from your memory when you say the names of your friends or family members, the place you live or the country you belong to you were accessing your memories.

When you say what is you can do or what is you cannot do what is your profession is, right you are expressing your memories. If you are stiff of your memories, your mental life and social life will literally fall apart and which is what happens in case of memory related brain disorders like Alzheimer's. As far as brain is concerned, memory is a trace left by ongoing neural activity in the brain. So, when you write on water, the ripples that you create disappear very soon, but when you write with a pencil on a paper you leave a trace you are creating a memory.

So, when sensory information that you receive from the world, flows through various path ways of the brain, it leaves traces in the brain these traces are in various forms they could be in terms of electrical activity of the neurons or in terms of the strength of this synapses or in terms of neurochemistry and so on. There is no centralized memory organ in the brain, there is no cerebral hard disk brains memories are distributed and curreted, by an extensive brain network. We do not know the full story of how and where and at what form memories are stored in the brain that is still a work in progress. In this lecture we will visit some basic ideas of how memories are organized and processed by the brain.

(Refer Slide Time: 01:49)

Brain stimulation and recalled memories



- Wilder Penfield (1891-1976) performed earliest experiments on 'brain mapping by electrical stimulation.'



So, let us start our story with Wilder Penfield and his experiments on brain stimulation. So, we have seen in earlier lecture how Wilder Penfield performed his stimulation experiments on patients and you know constructed some of the earliest brain maps. So, in wilder penfields patients were often epileptic you know patients. So, one of the surgery surgical treatments for this kinds of patients is to remove this source of the epileptic activity in the temporal lobe.

So, when during this surgery, in order to find the exact locus of the you know elliptic activity, a penfield used to stimulate the temporal lobe and find stimulation of which part produces a epileptic activity you know in the patients.

(Refer Slide Time: 02:34)

Stimulation produced Hallucinations

- Stimulation of temporal lobe produced visual, auditory, olfactory hallucinations and memory recall
- Subject begins to relive "an earlier period of time in the greatest detail."
- The neurological sequence corresponds to the forward direction of the original experience, never in backward direction



So, well he was doing that, he found that a stimulation of temporal lobe produce visual auditory all and factory hallucinations and memory recall right. For example, subject felt that they were able to recall experiences from long past and so, often this exceeds a memories or recall memories had great detail. And also neurological sequence are corresponds to the forward direction of the original experience and never in backward direction.

(Refer Slide Time: 03:09)

Auditory Hallucination

- In one of Penfield's patients:
 - Stimulation site: **Right superior temporal gyrus**
 - **heard a specific music selection** being performed by an orchestra
 - The sound was so clear that the patient believed that there was a phonograph in the operating room.
 - As the **same point** was restimulated the **music began at the same spot in time** where it had previously begun.



Suppose you recall a song which always in the original sequence of the song never in the backward direction. And in one case when the Penfield's patients recalled you know auditory had a auditory hallucination, the stimulation site was right superior temporal gyrus, the patients heard a specific music selection which was being performed by an orchestra.

This song was so, clear that the patient thought there is actually a phonograph playing in the operating room. At the same point when it was stimulated again the music began at the same spot in the time so; that means, stimulation at given point always produced right music from the same point in this sequence.

(Refer Slide Time: 03:46)

A place of
"psychical responses"

- The areas which can be stimulated to produce the record of the past all lie in the **temporal cortex**. Stimulating others areas of the brain had no effect.



So, what is interesting is only stimulation of temporal lobe produces kind of memory record. So, the areas that can which it can be stimulated to produce record of the past, all lie in the temporal cortex. For example, you can stimulated the primary visual cortex.

It will produce flashes of you know light right or if you stimulate the somatosensory cortex, it will produce in experience of being touched, but its not it does not give a rise to a recall of past memories. Whereas, stimulation of the temporal cortex very often right produce the recall of past memories. So, in fact, a Penfield called this area a place of psychical responses. So, that is what he called the temporal lobe.

(Refer Slide Time: 04:28)

The search for the Engram*

- Engram is a trace of memory in the brain, a term coined by Richard Semon in 1921.
- Specific pools of neurons in the brain are thought to serve as physical substrates for specific memories
- Karl Lashley began his search for the elusive Engrams in 1920s.



NPTEL

So, later on is the search for Engram right or the engram is basically a word coined by Richard Semon to denote memory to denote the neural substrates of memory. So, people thought that memory is are stored in specific locations in the brain and in say in terms of specific process of neurons. So, they thought if they can stimulate those neurons and they could recall memories Karl Lashley launched his search for Engram right he did interesting certain experiments, to find out these neural substrates of Engrams and this studies have began sometime around the 1920s.

(Refer Slide Time: 05:04)

The search for Engrams begins

- Approach: effect of brain lesions on maze learning in rats
- Observation: found that severity of the learning deficit produced by brain lesions depend on the extent of the damage, not on its precise location.
- Conclusion: Learning doesn't have a special locus in the brain.

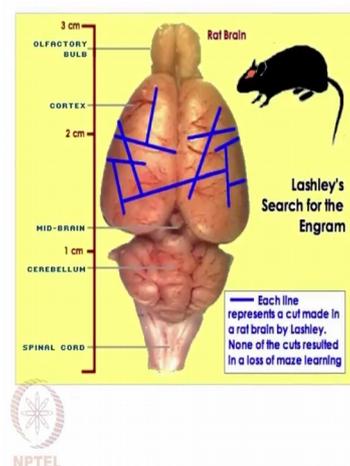
- Search for Engrams was inconclusive!



NPTEL

So, his approach was like this. So, he made he took a experimental animals. So, specifically he took a rats and rats are where maze learning as we all know. So, if you leave rat in a complex maze, it would under own the maze for sometime become familiar with the maze and then learn how to escape from the maze so; that means, it is its memorizing the maze, the maze is layout in its brain and then using that memory right to escape from the maze. So, Lashley wanted to study what exactly is substrates for this memory in the brain of the of the animal.

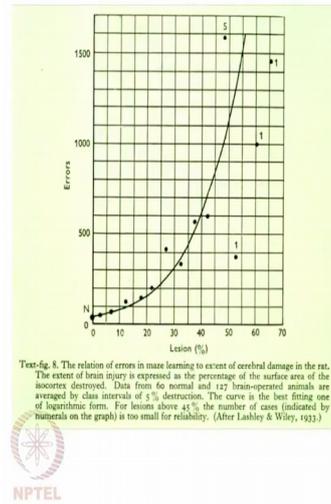
(Refer Slide Time: 05:35)



So, to this end he would actually make cuts in the in the in the rats brain, this cuts are shown this image as you know straight blue lines, and he wanted to know whether there is a correlation between these location of the cut and the deficit in the memory recall; that means, its efficiency in the performance of escaping from the maze.

So, what you found is, the learning or memory does not have a specific locus in the brain, it only dependent on the extent of damage and not on precised location of the damage. So, so for example, he control he length of the cut, and he also vary at the location of the cut and what lashley found is that the location does not matter much and only the length of the cut matter. So in fact, he made graphs like this.

(Refer Slide Time: 06:19)



So, in this graph the x axis represents the extent of this lesion and the y axis represents a errors made by the rat in escaping from the maze. And you see that there is a almost a kind of monotonic its not linear, but a monotonic relationships between the lesion and the errors. And in fact, as a lesion got bigger and bigger then number of data points on this graph are fewer.

Probably because as a region got bigger the animals did not survive this, the kind of a drastic surgery. So, the from this experiments; so, lastly concluded that there is no locus for memory and memories are distributed all over the place.

(Refer Slide Time: 06:58)

Lashley's Principles for the Nervous System

- **Rejected the notion of localized learning.** Rather he proposed **two principles**:
- **Equipotentiality:** All parts of the cortex contribute equally to learning; one part can substitute for another part.
- **Mass Action:** The cortex works as a whole; performance improves when more of the cortex is involved.

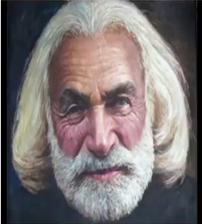


And out of this studies he proposed two principles how brain operates first one is called equipotentiality; that means, all parts of cortex contribute equally to learning. So, and one part can in can substitute for the other and the other principle is law of mass action, that is cortex works as a whole performance improves when more of the cortex is involved. So, its goes to totally against the classical localization principle that had helds way in neuroscience for a for many centuries.

(Refer Slide Time: 07:27)

Memories are holograms

- Karl Pribram one of Lashley's students continued the search for the engrams and...



Karl Pribram (1919-2015)



...saw the analogy between **Memories** and **Holograms**

So, then these ideas for taking up further by one of Lashley's students who is Karl pribram and he continued his search, and his search began by a perception of the analogy between memories and holograms.

(Refer Slide Time: 07:40)

Holograms

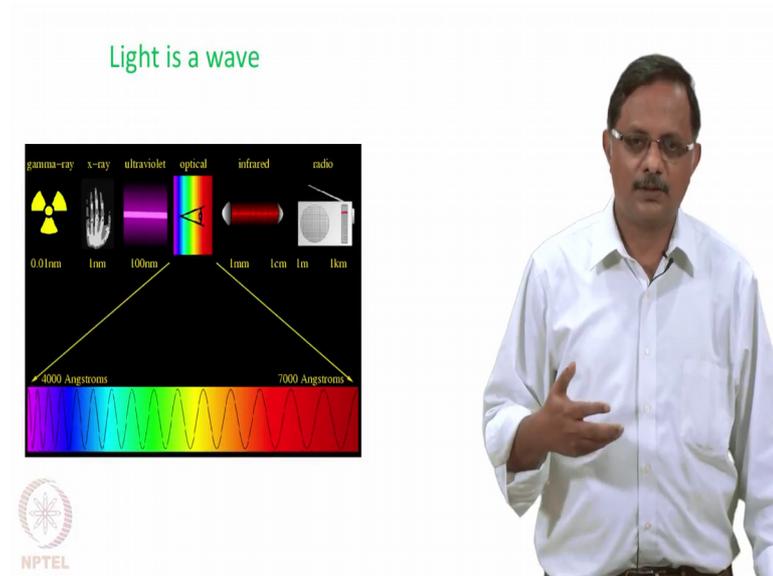
- Holograms are 3D-like displays of objects
 - Unlike images which look the same as view varies...
A hologram varies with varying view
- They are based on the principle of Interference of Light



 NPTEL

So, holograms are three d like displays of objects and we see them very often in popular media in movies in especially in science fiction or futuristic movies; unlike an image right hologram is slightly different from an image and an image no matter which way you look at it from it looks the same whereas, a hologram when you go around the hologram and look at it from different van touch points, hologram look slightly different. So, it gives a kind of a illusion of a three dimensional object and the genesis of holograms are based on the principle of in interference of light.

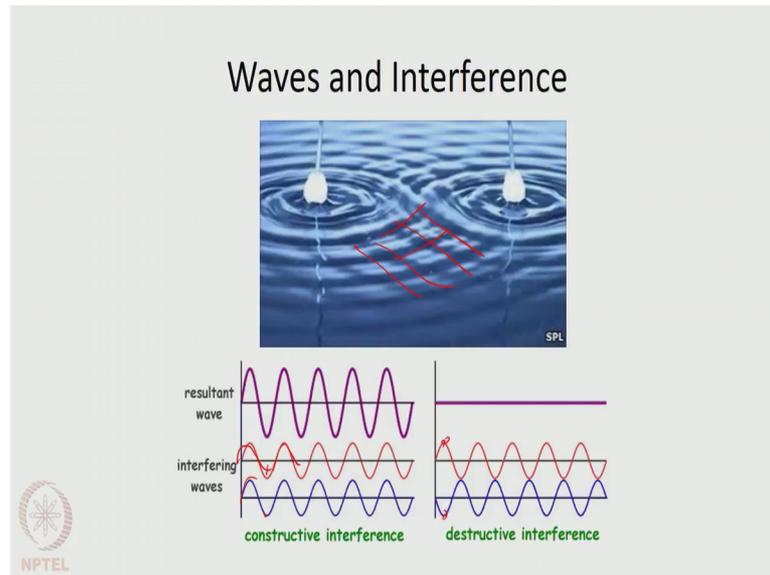
(Refer Slide Time: 08:12)



So, we all know that light is a wave like any wave, it has an amplitude and a frequency, and like a propagating wave it has also a wave length and it is also has a face. So, now, we know that. So, in this picture you see the wave lengths of different lights different types of radiation.

So, the visible light has is wrestle between the wave lengths of about 100 nanometers to 1 millimeter. So, beyond that if you look at longer wave length have infrared between about one millimeter to one centimeter and beyond that radio waves go from meters to kilometers in wave length. If you go below visible light in the range of about 100 nanometers you have the ultra violet radiation and if you go to x rays which has much lesser wave length we have about 1 nanometer wave length for x rays and if you go below that gamma radiation which is a very strong radiation, higher degree radiation has wave length of about 0.01 nanometers.

(Refer Slide Time: 09:11)



We all know that if you drop a pebble in a pond of water you know it creates ripples, from the point where the pebble hits the water. So, these ripples spread in all in a circular fashion around the point of contact. So, in this picture you see two such rods emerged in water and making contact with water and these two rods are vibrating, creating a ripple at two different points on the water surface.

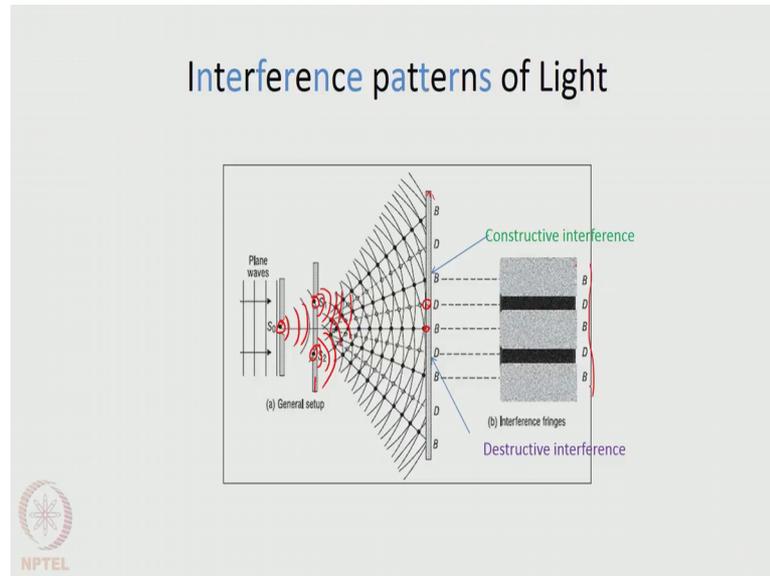
So, you have this circularly expanding waves from each of these points of disturbance, and these waves have the spread they interfere and mix and mingle at some point and you can see that. So, where they mingle the two circles meet and form a kind of a grid-like pattern right and these are called interference patterns.

So, basically there are two kinds of interference. So, in the bottom part of the figure you see that right you have two waves both have the same phase; that means, both hit their maximum points which is called the crest at the same point at the same position, and hit their minimum point which is called a trough right at the same point and so, the crests and troughs occur at the same position same point.

So, therefore, they interfere constructively with each other producing a much larger wave. So, but on the other hand if the crests and troughs do not align, and where there is the crest of one wave if we have the trough of another wave, to this such waves tend to cancel out each other and such interference is called destructive interference.

So, in here interference pattern like this, you see that at some point the waves interfere constructively where you have a high amplitude of the wave, at some points the waves interfere destructively, where you have a low amplitude of the wave.

(Refer Slide Time: 10:55)



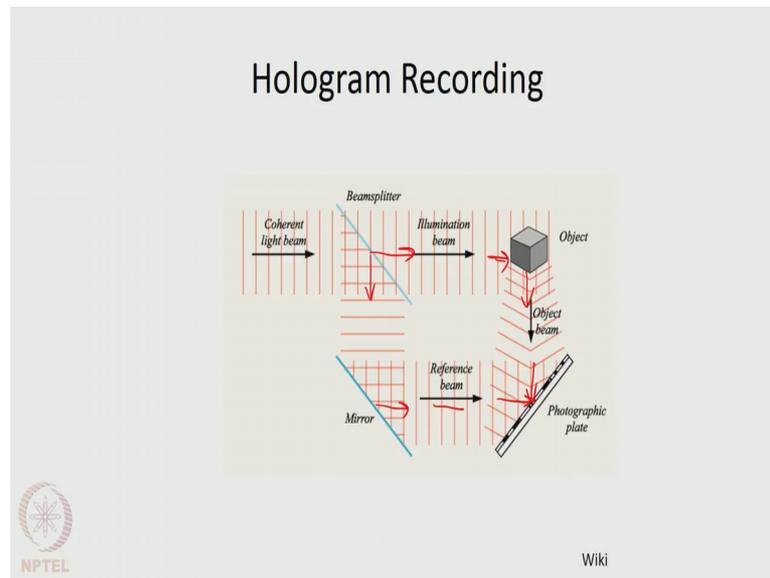
Since light is also a wave, right light waves also interfere and produce this kind of a constructive and destructive interference patterns. So, in this picture in fact, this was a classic experiment performed by Thomas Young you know several centuries ago which was a very definitive experiment to prove the wave nature of light.

So, in this picture you have a plane wave of light hitting the wall and the left corner right and this wall has a slit and so, light which this slit acts like a point source of light, and from this point onwards lights spreads as circular waves around it, and this circularly expanding wave hits the second barrier, and this barrier has two slits in it and. So, the circularly expanding wave hits two slits and these each of this slits now serves as a as a points source.

And you have circularly expanding waves originating from both of this slits, and these waves now intersect each other they interfere with each other forming a grid like pattern. And when these interfering waves hit these screen at the back of this [ex/experiments] experiments right at some points on the screen the interfere constructively forming a bright spot at some point the interfere destructively forming a dark spot. So, what you see on this screen is a interesting black and white pattern which is what you see here on

the on the right most edge. Now this kinds of a black and white patterns can be photographed, if you put a photographic screen right at the at the far end, then you can take a picture of this interference pattern.

(Refer Slide Time: 12:31)

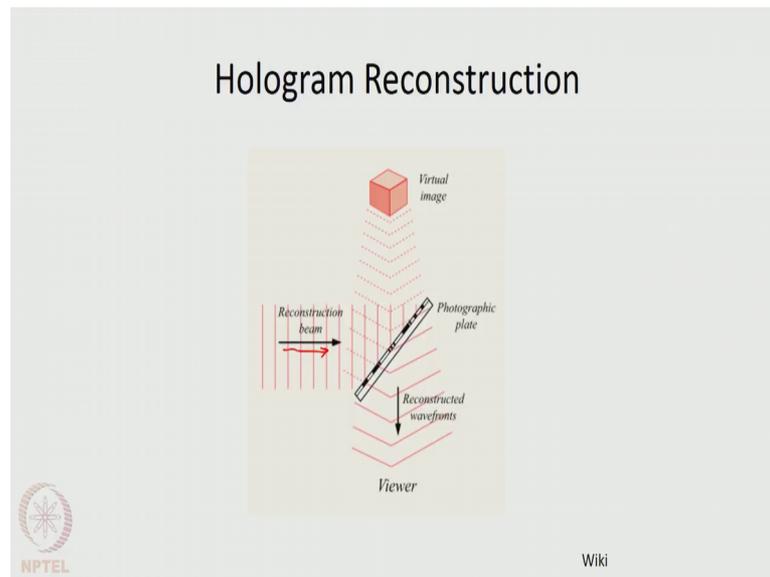


Now, this is what is used in creating a hologram. So, in a hologram we have a coherent light beam that is a light beam which has only a single frequency and all the waves are in phase. So, usually a laser has this kind of a property. Now a coherent beam hits a beam splitter.

So, at the beam splitter some waves proceed further on you know crossing the beam in some wave reflected by the beam and get deflected right and move downwards; and this kind of a deflected beam hits a mirror and then again is deflected by 90 degrees and moves to the right and this is called the reference beam.

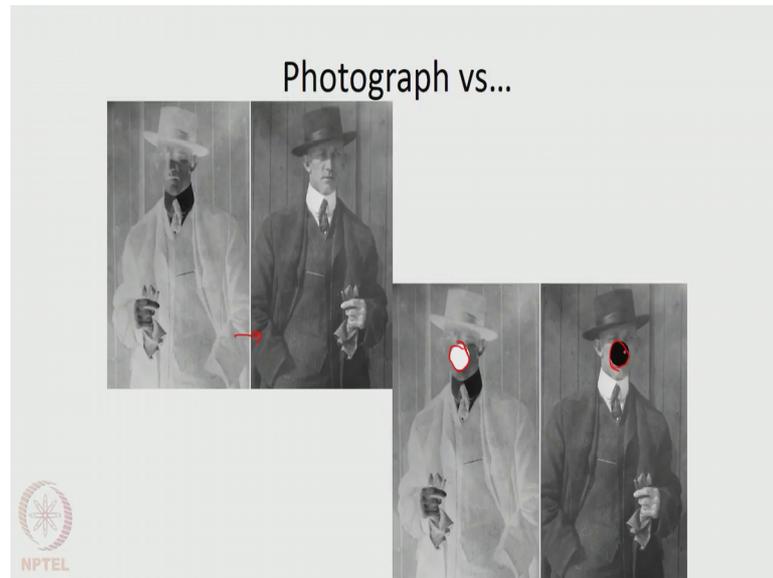
And then on the other side you have the beam which goes through the beam splitter goes straight and falls on the object, and once it hits the object it gets reflected by the object and the reflected waves from the object again come downwards. So, at the photographic plate the beam from the object which is called the object beam and the beam from the reference beam which comes directly from the source right which is called the reference beam, both these beams interfere forming an interference pattern and this interference pattern is captured on a photographic plate.

(Refer Slide Time: 13:43)



So, this photographic plate is used then to recreate the same kind of a beam that is created when the light falls on the object and gets reflected by the object. So, to recreate that you use a another reconstruction beams, this also coherent beam of light and then you have the photographic plate, which we created in the previous situation and when this waves pass through photographic plate, the reconstructive wave fronts see what you see you seen by the viewer from the other side, and what actually looks to the viewer looks like a three dimensional object the original three dimensional reconstruction of the object. So, the when you reconstruct a hologram using this kind of a interference pattern printed on a on a photographic plate right.

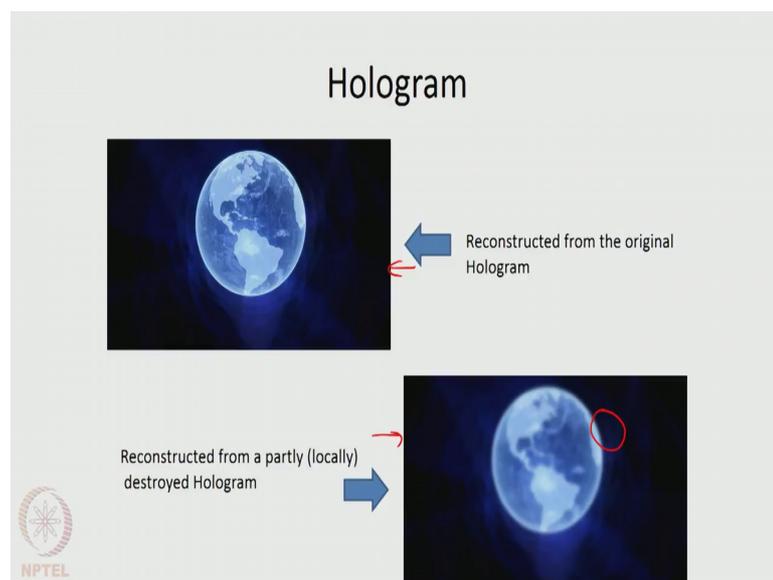
(Refer Slide Time: 14:25)



Their experience quite different from what you get when you I mean develop a negative and create a positive. So, when you develop a negative suppose you damage a small part of the negative.

So, here you have negative and its corresponding positive, you damage a small part of a negative. I mean you develop it in the positive right you see a small local damage corresponding to the part where you have damaged in the negative.

(Refer Slide Time: 14:49)



But when you reconstruct a hologram. So, if you take a hologram the interference pattern as it is printed on a photographic plate, you take the photographic plate tear up a small part of it right and reconstruct hologram. So, here on top you see the hologram from the original complete and intact interference pattern.

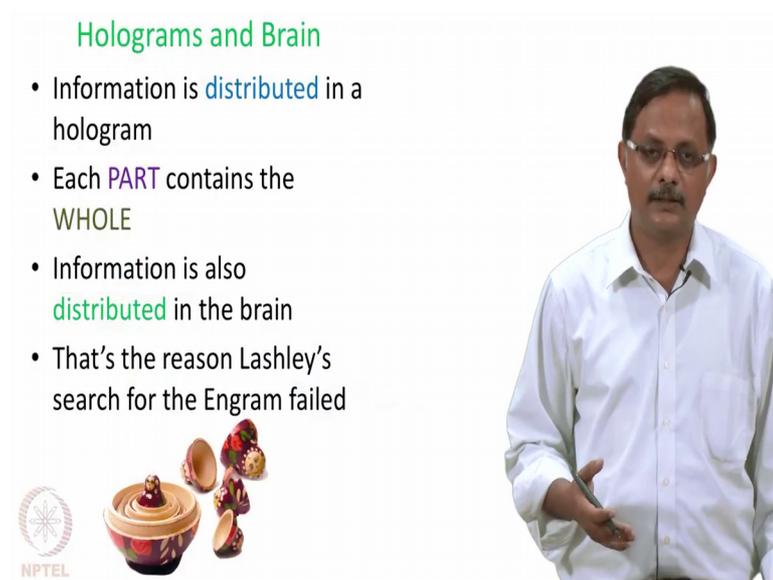
So, and in the bottom you see you tear you see a reconstruction from a kind of a damaged hologram pattern, which locally damaged. But in the in the reconstructed image in the hologram we do not see a local deficit you know like what you have seen in case of a photograph you see that the whole image is there, but the whole image is certainly degraded ok. So, it is as a information distributed all over the place inside the holographic film.

So, that is a main difference between holograms and image and that is where holograms can be compared to the memories was stored in the brain that is what pribram thought.

(Refer Slide Time: 15:43)

Holograms and Brain

- Information is **distributed** in a hologram
- Each **PART** contains the **WHOLE**
- Information is also **distributed** in the brain
- That's the reason Lashley's search for the Engram failed



The slide features a man in a white shirt on the right side, gesturing as if speaking. Below the text on the left, there is an illustration of a Russian doll (matryoshka) and the NPTEL logo.

So, information distributed in a in a hologram, it is also each part of the picture is stored in the stored everywhere in the in the holographic plate. So, in early information its also distributed in the brain. So, that is the reason perhaps Lashley's search for the Engram failed because he was not show any one path and there everywhere. Something like what you find in a Russian doll.

(Refer Slide Time: 16:03)

Holonomic Brain Theory

- Proposed by Karl Pribram
- Idea: Brain networks transform the input signals into the frequency domain and construct hologram-like representations of the world



So, based on these kind of analogy this kinds of insides pribram proposed a homonymic brain theory, basically it proposes that brain network stands from input signal into the frequency domain and then construct hologram like representation of the world.

(Refer Slide Time: 16:19)

- Brain dynamics consists of a variety of oscillations – α , β , γ , δ etc.
- There are frequency-tuned neurons in auditory, visual and somatosensory systems
- Fourier-like transforms are known to occur in the auditory cortices
- Yet a far simpler model is possible...



So, there is a which was very elaborated soph[isticated] mathematically quite sophisticated theory that was proposed by pribram and in fact, there is lot of evidence from neuroscience from neurobiology to support that kind of a theory. For example, right to create pribrams theory, we need you need to operate in the frequency domain.

You need to deal with oscillations and we know the brain dynamic consists of a variety of oscillation. So, you know in various frequency bands like you know alpha beta gamma delta all these frequency waves are seen in the electroencephalogram or EEG. So, there are this frequency tune neurons in the auditory visual and somatosensory systems. So, in auditory you find neuron which are tuned to certain tones, pure tones in the visual cortex you will find neuron which are tune to spatial frequencies.

So, if you show a periodic wave like pattern like visual pattern, you will find neurons in a primary visual cortex which are tune to it and similarly in the somatosensory cortex you will find neurons which are tune to a frequency, that is if you apply a vibration right your skin at certain frequency right you will find certain neurons responding to it. Now Fourier like transforms are known to occur in the auditory cortices for example, in there is lot of work done on bats auditory system right here. So, the input frequencies are transformed to some kind of Fourier transforms.

So, that you know on a strip of cortex. So, that neurons at response to low frequencies are at one end of the strip and neurons are respond to high frequencies or the other end of the strip. So, that is like a Fourier transform that is going constructed in the auditory cortex. So, it is in principle possible to construct and to map they have holographic brain theory, on to you know real by neurobiological data.

(Refer Slide Time: 18:05)

The Hopfield neural network

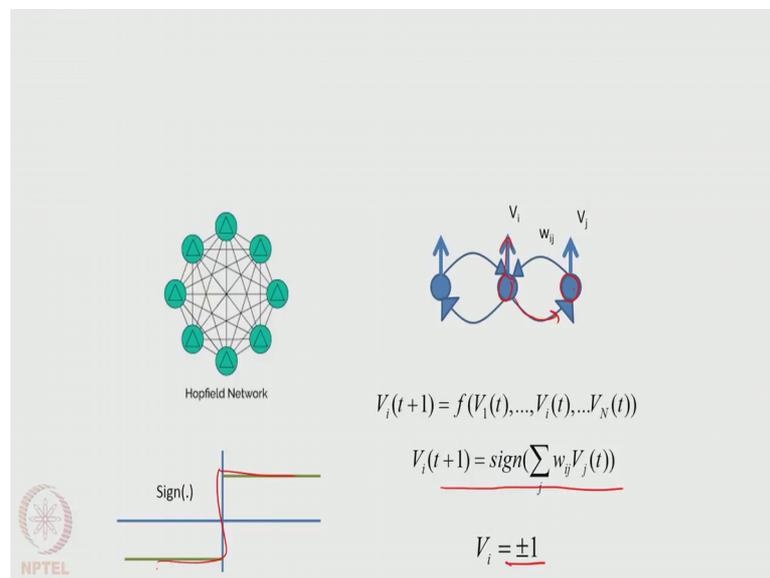
- Proposed by John Hopfield (1982, 1984)
- A network of McCulloch-Pitts neurons
- Each neuron is connected to every other neuron
- Can serve as a memory
 - Patterns can be stored and retrieved
 - Information is stored in a distributed fashion in the weights



But actually let us begin with a much simpler model and in fact, let us look at a much simpler model, which also give some of the same properties and you do not need to operate in the frequency domain and talk in terms of oscillations. So, this model was proposed by John Hopfield in a series of papers 1982 and 1984 interestingly this model works constructed using the same McCulloch Pitts neurons, which we visited in some of the earlier lectures. So, in this network each neurons [cons/connected] connected to every other neuron. So, unlike some of the feed forward network you that you have seen some of the earlier lectures, where neurons always connected in a feed forward fashions.

So, you had a series of layers and lower layer projected to higher layers and so on so forth you never had a loop anywhere in the network. So, neurons never fed back to neurons from where they receive inputs. So, in Hopfield network every neurons connect every other neurons this kinds of networks it was shown can serve as a memory can store patterns and retrieve them and like in a memory. And its also been shown the information this networks restored in a distributed fashion in the weights.

(Refer Slide Time: 19:14)



So, here is simple schematic of the Hopfield network. So, see a network diagram on the left and. So, you see a more detailed picture where you have you know three neurons. So, the i th neuron and j th neuron here. So, i th neuron each neuron has a state v_i . So, V_i is a state of the i th neuron and its states are binary like in the McCulloch Pitts neuron.

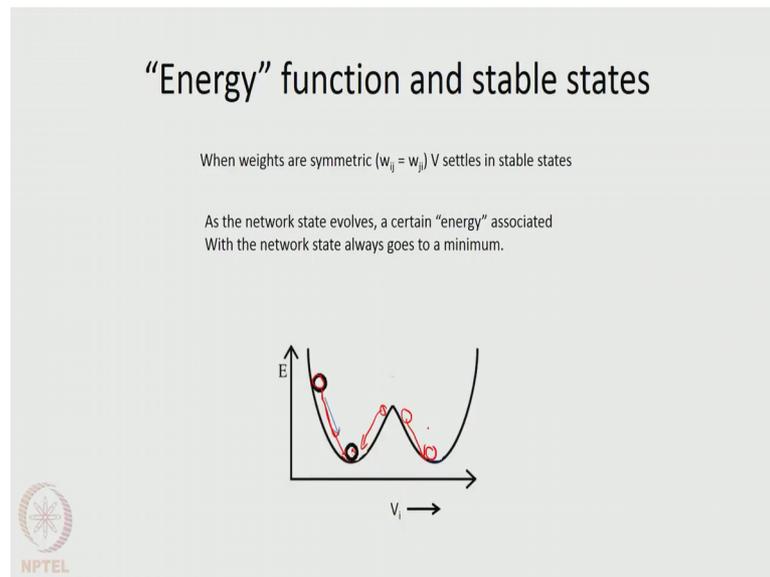
But only difference is the states are typically plus or minus 1 in Hopfield network whereas, in McCulloch Pitts these states are 1 and 0, but both are equal and then you can work the theory completely in terms of 1 0 types state or you know 0 1 plus or minus 1 types states both are equivalent. So, the each neuron receives input from every other neuron. So, and so neuron i receives input from neuron j and neuron j in turn receives input from neuron i .

So, there are you know the network is full of loops. So, what does; that means, is unlike in the static feed forward networks, where when you give an input to the input layer and the output of that goes to hidden layer output of that goes to the output layer all this happens instantaneously there is no time delay involved in this kind of a transmission. Whereas, in case of Hopfield network because of the loops right when V_i is receiving input from V_j , V_j receiving input from V_i there is a kind of a conflict in terms of sequence which one which input goes to which neuron first.

So, therefore, to resolve that they only way to explain that is when you are looking at V_i in neuron, you get it gets inputs from all the states of all the other neurons at the current time and then updates its state and then the state becomes a next state. So, therefore, we have to automatically bring in the time factor in describing the information exchange in this kind of a network. So, therefore, V_i at this t plus one next state is a function of this states of all the neurons at the current time that is t and this.

So, this is how the McCulloch Pitts neuron model is used in the Hopfield neuron. So, you take a linear sum of all the neuron states weighted through the weights w_{ij} and pass it through a signum function which is basically this sign function and the output is now given as the next state of the i th neuron and the same thing is done repeatedly for all the neurons and again and again for a entire network.

(Refer Slide Time: 21:37)

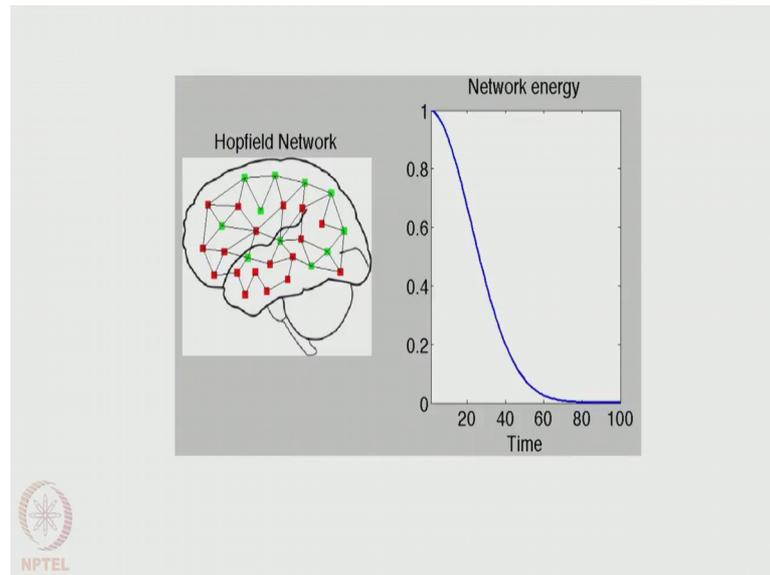


So, it was shown that this network is associated with a function called the energy function there is nothing to do with a physical energy.

But it is motivated by some of the physical systems. So, you know that in lot of physical systems energy goes to minimum. So, for example, when you when you drop an object right the potential energy of the object goes to minimum and its so, object keep moving until the potential goes to minimum. So, similarly in case of Hopfield network there is a certain energy function its hypothetical kind of numerical energy function, which goes to minimum as the states of the network keep changing driven by this kinds of update rules. So, so in this schematic you can see the state of the network going down to a minimum and settling there.

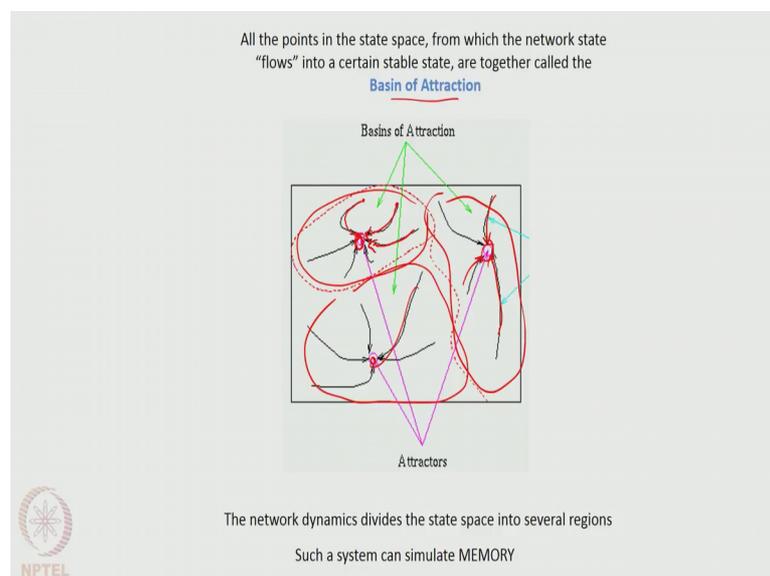
So, then this energy function typical as multiple minimum. So, this if this networks starts from here, it will rolled on and settled here it starts from here it rolled on and settled here if its starts from here roll rolled on and settled here and so on so forth. So, they can be multiple minima of this energy function.

(Refer Slide Time: 22:35)



So, this can be seen in this kind of a simple simulation. So, basically this network is kind of superimposed on a cartoon of a brain just to make it look interesting. So, the flashing points are all imagine them to be neurons and green stands for plus 1 and red stands for minus 1. So, as the network evolves governed by this update rules, this state keeps changing on the right you see this curve which represents energy and energy is going down going to minimum value now it is flattening over at a minimum value. As it approaches the minimum you can see that the state also stops evolving and its settles down in the stable state.

(Refer Slide Time: 23:13)

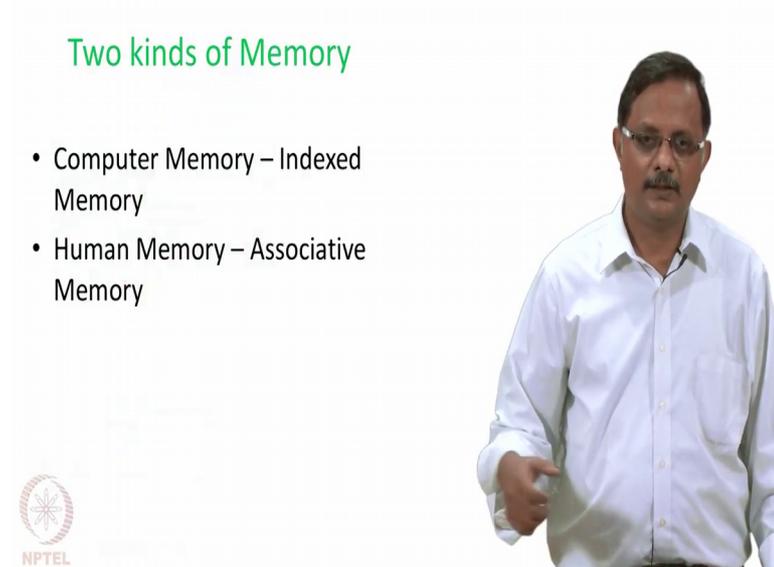


Now, they can be seen as stable states, because the energy function has minima as minima. So, what happens in this kind of dynamic situation is if you start from certain initial conditions, you will evolve towards the minima of the energy and go to a stable state, but.

So, in this example in this cartoon picture, you can start from many of these initial conditions and go to the same minimum. So, similarly, if you start from other initial conditions, many of these initial conditions will go to a same new, but this is a different minimum than what we saw before. So, all the initial conditions which take you to a given minimum are together called the basin of attraction, right, corresponding to that minimum. So, therefore, the entire state space of the network in this case is divided into multiple basins of attraction. We have one here, we have one here, and then another here, and so on, so forth.

So, the whole space of the state space of the network is given to these multiple basins of attraction. If you start from anywhere within a given basin of attraction, you will roll down and settle at the corresponding minimum for that basin of attraction. And basically, these basins of attraction, these stable states of the network dynamics, are thought to be memories. Let us see how that works. So, now, before explaining how Hopfield networks work as a memory and comparing with human memory.

(Refer Slide Time: 24:36)



The slide features a title 'Two kinds of Memory' in green text at the top. Below the title is a bulleted list: '• Computer Memory – Indexed Memory' and '• Human Memory – Associative Memory'. In the bottom left corner, there is a circular logo with a star-like pattern and the text 'NPTEL' underneath. On the right side of the slide, a man in a white shirt and glasses is standing and gesturing with his right hand.

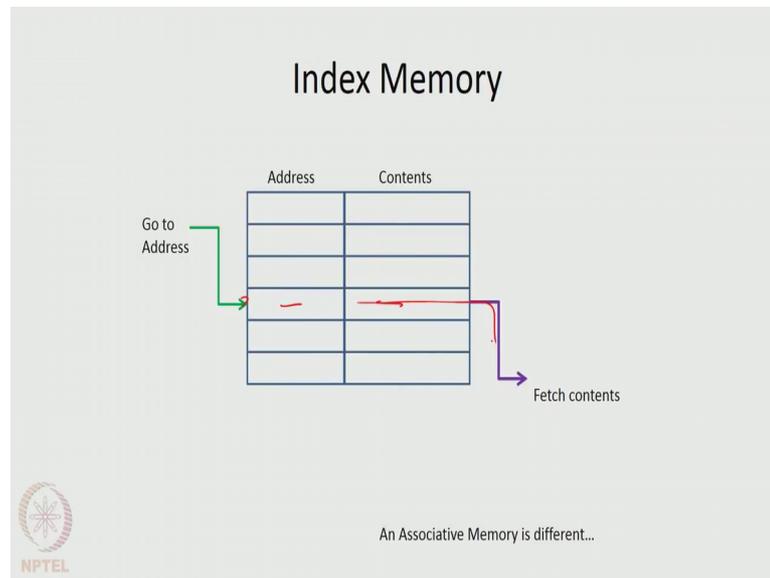
Two kinds of Memory

- Computer Memory – Indexed Memory
- Human Memory – Associative Memory

NPTEL

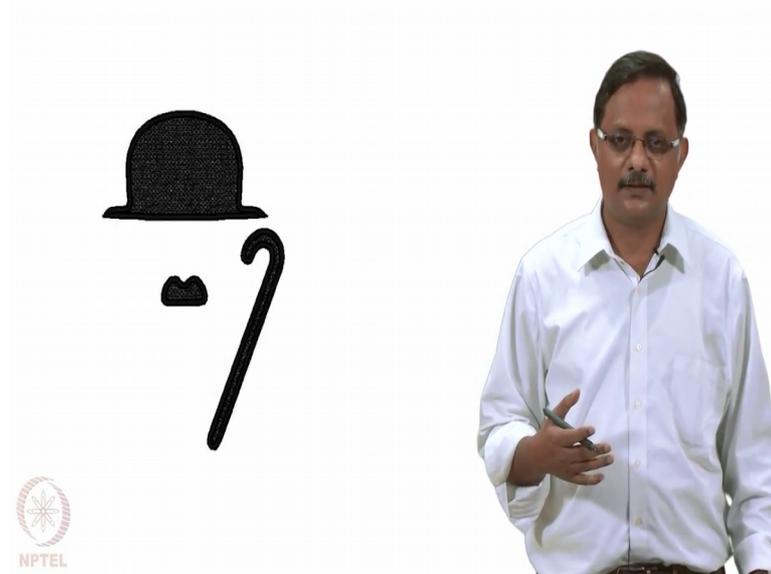
We need to state that there are two kinds of memory the kind of memory that we have we seen in a computer and human memory and we should remember that these are very different kinds of memories. So, let us look at a computer memory.

(Refer Slide Time: 24:49)



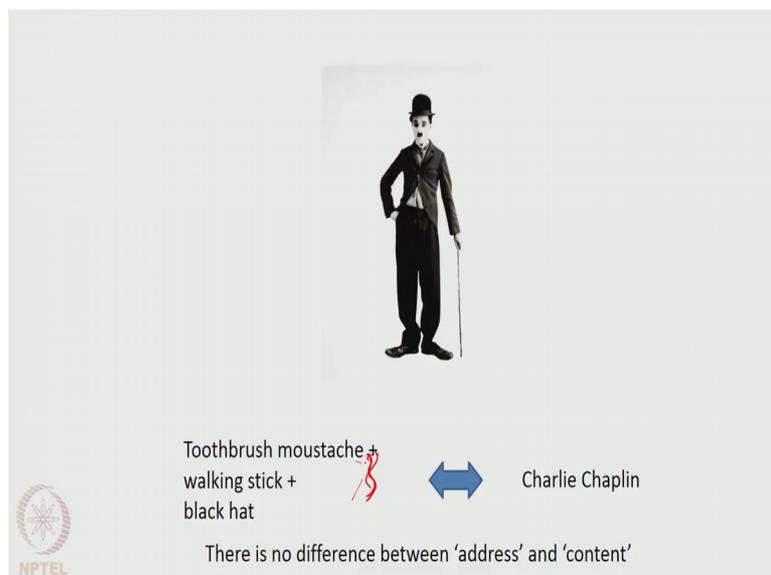
So, in case of computer memory the information is stored in a long kind of a row, where at each slot there is an address and then if you go to the address you retrieve the contest. So, if you want to retrieve the content which is here, you need to know what is the address here and go to that location and then the retrieve the content. But that is not how human memory works to understand that let us look at a small example.

(Refer Slide Time: 25:14)



So, I am you are looking at a interesting sort of black like a block in this screen right can you associate with some person right this is picture is not sufficient. So, let me give you one more clue, what about this what does the what are these objects remind you all? Right If it is not clear even now let us look at this well now it is very clear it is the Charlie Chaplin.

(Refer Slide Time: 25:37)

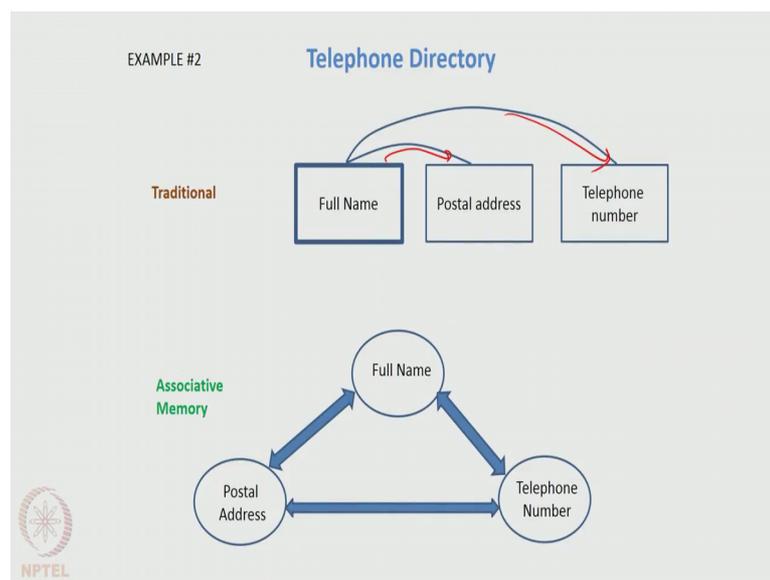


So, thing is the combination of a toothbrush moustache and a walking stick and a black hat right can be easily associated with Charlie Chaplin or vice versa. If you think of

Charlie Chaplin you think of all this objects and sometimes even a subset of all this objects may be enough to associate with the person Charlie Chaplin. So, in this kind of association. So, it so, when you do this you are recalling from memory.

So, when you look at these objects you are recalling the memory of the person Charlie Chaplin or when you look at the name Charlie Chaplin, you are recalling the corresponding attributes of that person. So, in this case there is no difference between the address and the content right there is no such distinction. So, if you give the properties or attributes as some kind of a address right you are recalling the name Charlie Chaplin as a content or other way around, if you are given Charlie Chaplin as a address like and then you are recalling here attributes has some kind of a contents. So, there is no distinction between address and content in the way we memorize things.

(Refer Slide Time: 26:37)



Let us look at one more example. So, let us look at the traditional telephone directory. So, we use to have this yellow fat books right which we where we used to look up phone numbers of people. And if you remember the way these phone numbers are organized in a telephone directory is you have the last name of the person, then the first name then the postal address of the person and the telephone number.

So, if you; so, if you need to find out the telephone number you need to know the full name, the last name and the first name and once you have that from that you can find the postal address and you can also find out the telephone number. But ideally in a in an

ideal telephone directory what should you like to have is, every possible association between this three items right. So, given the name you should be able to find the address and the telephone number or given the telephone number, you should also be able to find out are to whom this number belongs right or where does this person stay, what is the address right. So, that is in associative memory. So, in this kind of a set of there is no distinction between the address and the content.

(Refer Slide Time: 27:41)

Human Memory

Is an
Associative Memory

Hopfield network can
Model an
Associative Memory



So, therefore, human memories more like an associative memory, and Hopfield network can model an associative memory, and as we will see that very easily. So, if you want to use Hopfield network as a memory and then in if you want to you need to figure out how to store information in the in Hopfield network, and that is done by a simple rule.

(Refer Slide Time: 27:58)

Pattern storage

If S is the pattern to be stored,
Weights are calculated as:

$$w_{ij} = S_i S_j$$

If multiple patterns S^p are to be stored,
Weights are stored as:

$$w_{ij} = \sum_p S_i^p S_j^p$$


Which is given here and this rule is called Hebb's rule and we will talk about that in a moment. So, basically if you want to store a pattern, which has a bunch of bits these are like a bunch of s plus ones and minus ones. Then given the pattern you basically calculate the weights as; that means, the pattern is stored in the form of weights. So, using this rule. So, w_{ij} that is a connection taking from i th neuron to j th neuron is nothing, but the product of the i th bit and the j th bits.

So, s_i times s_j if you want to store multiple patterns doubly slightly extended like this right. So, s the p corresponds to the index of the pattern. So, s_{pi} is the i th bit of the p th pattern and s_{pj} is j th bit of p th pattern. So, it takes similar products as what we have done before for a single pattern and takes as products for all the patterns and add them up, and that that becomes here the value of the ij th element or ij th bit.

(Refer Slide Time: 28:58)



We just did that and we stored three images in a Hopfield network and then we try to retrieve this images by giving a small part of the images a q right and we ask as. So, we give this written part of this image as a q and make the network retrieve it right it will. So, quickly retrieve the whole that part is given was from the Monalisa image and it is complete Monalisa image and you can give this part from any part of the whole image.

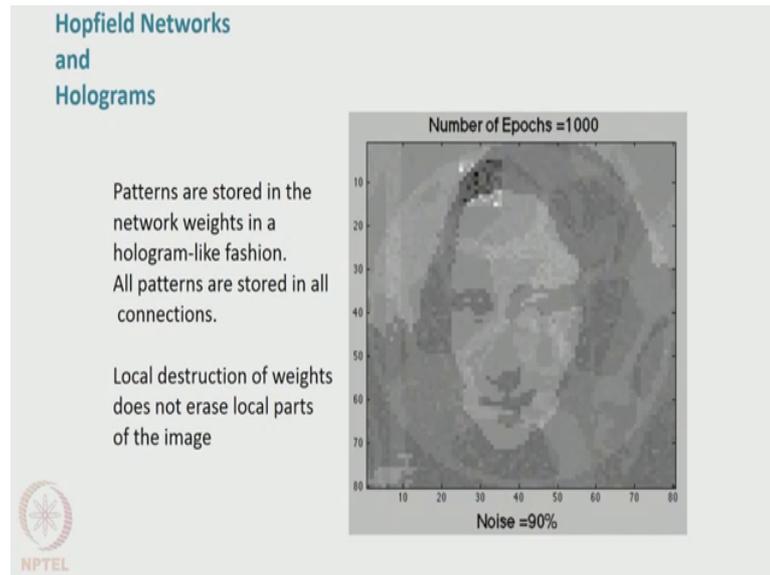
So, in the next stimulation right takes little while to retrieve entire image because here each neuron is updated you know individually one after another. So, now, you give a different part of the image. So, again it retrieves the full the full image. So, it is associating each part of the image with the whole ok. So, there is no distinction you do not have to give a specific part of the image there is nothing like an address for the full image.

(Refer Slide Time: 30:05)

**Hopfield Networks
and
Holograms**

Patterns are stored in the network weights in a hologram-like fashion. All patterns are stored in all connections.

Local destruction of weights does not erase local parts of the image



Number of Epochs =1000

Noise =90%

NPTEL

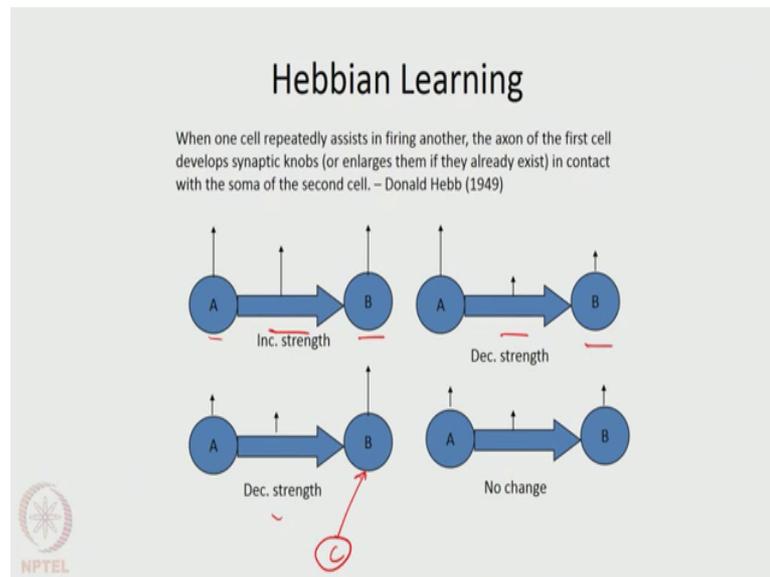
Now, just like an hologram right the information distributed all over the weights. In fact, the information about all three images are distributed in all the weights, and also damaged to some weights does not create a local damage to the network.

And damage is seen all over the image. And you see a general degradation a certain degradation over the entire image and also its quite robust to damage as you can see in this animation. So, in this case again we give a small local part of the image and deconstruct it. So, what you are seeing right now is a network in which ten percent of the weights are damaged.

So, we randomly set, 10 percent of the weights to 0 and reconstruct it its able to get the image pretty reliable. So, again in this case its 50 percent of the weights are set to 0, again its able to recall very accurately. In the third case we have damage 90 percent of the weight. So, here it is pulling up the image, but it is actually take much longer and what you see in this case is always recalling Monalisa image, in the background you see a faint picture of the ball and the little bit the third image.

Which is a picture of the lena right. So,. So, in this case recall is not perfect, but you can still see kind of a faint picture of the Monalisa image all though the 90 percent of the weights are restored.

(Refer Slide Time: 31:30)

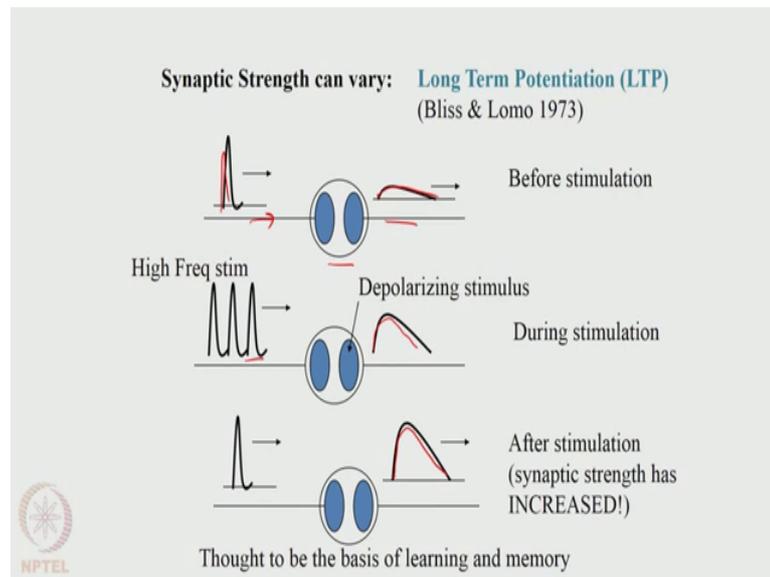


Now we have used a rule called Hebb's rule like I just mentioned to store images store patterns in the network and so in fact, Hebb when he used the running rule he actually derived it from mathematical considerations, but this rule was actually proposed by a psychologist long ago in 1949. So, by a person called Donald Hebb. So, Hebb in his book called organization of behavior, he proposes a rule wherein he says that. So, he basically speculates about what could be going on to the level of cells right when something is learnt by the brain. So, what he says is when two cells are interacting and if cell neuron A is active and it also succeeds in making B active right then the connection between the two neurons must be strengthened which is what you see in the first case.

So, A is active and B is also active and when that happens often the connection strength between A and B should be strengthened. So, in the second case A is active often, but B is inactive. So, or not very active then the connection strength between the two must be decreased and the third case A is not active, but B is active and this could happen because B could be receiving input from other neurons like C for example.

When this happens often then the connection strength between A and B gets weakened. In the fourth case is trivial I mean neither A nor B is active. So, there is no change in the weight. So, these what Hebb has proposed could be happening at the level of cells and connections in the brain and brain is learning something.

(Refer Slide Time: 33:02)



Now, there is evidence that kind of a thing would be is actually happening in the brain and the some of his work the earliest work of this kind came from the work of Bliss and Lomo 1973. So, where they worked on hippocampus synapses and. So, in this experiment you have a synapse and there is an action potential which comes on the pre synaptic side when it hits the synapse.

So, it produces a post synaptic potential in this case it is a n p s p right as you can see a positive wave here. So, before stimulation, an action potential on the pre synaptic site produces small p s p on a post synaptic side. So, then they give a high frequency volley of impulses they pass this you know high frequency current pulses to the pre synaptic side and in response to that on the post synaptic side you can produce you can achieve a large amplitude e p s p post synaptic potential while stimulation is going on.

Then say in this goes on for about 10-15 minutes and after that again when you give a single action potential action potential on the pre synaptic side in response to this if it produces a much larger p s p then what you it produce in the beginning to the same input; that means, the. So, the synapse is producing a much larger p s p after the stimulation then before the stimulation, for the same input that is for the same single action potential; that means, in effect the strength of the synapse has increased because of the simulation ok.

So in fact, in this case when the high frequency volley of inputs are given to the pre synaptic side, simultaneously on the post synaptic side you pass a positive current thereby depolarizing the post synaptic membrane. So, activation of the pre synaptic side means giving high frequency volley of inputs, and activation of the post synaptic side means a giving a depolarizing current or a positive current.

When both of them are given simultaneously in the strength of synapse goes up and this phenomena is called long term potentiation which corresponds it corresponds to among the Hebb's cases it corresponds to the first case. So, a (Refer Time: 35:12) is also observed in this same experiment. So, like before you see you know you have synapse, where even action potential and the produce some kind of a p s p. Now you give a high frequency input on a pre synaptic side, but give hyperpolarizing stimulus that is you give a negative current to the post synaptic side. So, then because of this there is very little response on the post synaptic side because to hyperpolarizing membrane.

Then after doing this kind of a pairing for a few tens of minutes right if you give the single action potential on the pre synaptic side, the response to the single action potential on the post synaptic side is much lesser and what it was before the stimulation. So, in this case effectively the synapse has become weaker and this phenomena is called a long term depression and specifically this is this corresponds to the second case of you know Hebb's schematics.

So, in this case the A is active that is a because the high frequency inputs are given to the pre synaptic side, but the B is a inactive because the post synaptic membrane is hyperpolarized. So, when that happen the strength is weakened ok. So, we will stop this segment here.

(Refer Slide Time: 36:17)

Summary

- Memories in the brain are distributed
- Analogy with a hologram
- Human memories are associative
- Hopfield network as a simple Associative Memory
- Hebb's rule and LTP/LTD



To summarize we have pointed out right in the beginning that brains, memories in the brain are distributed there is nothing like a local site where memories are stored therefore, the search for Engram which was conducted by Lashley and others.

So, you know they have ended in conclusively and memories are like hologram, which is what pribram has you know has suggested because in hologram information is distributed all over the hologram right the each information about the whole is contending each part of the hologram and human memories are associative and they have this kind of a distributed property like hologram and Hopfield network which is model that we have described right is a model of a an associative memory and it uses Hebb's rule for pattern storage and retrieval.

And Hebb's rule is quite has lot of neurobiological support because there is even there is a early experiments on LTP and LTD which supports Hebb's rule. So, in the next segment of this lecture we will look at how memories are stored in hippocampus and what kind of a cellular and electro physiological machinery of hippocampus supports memory storage and retrieve.

Thank you.