Advanced Cognitive Processes
Dr. Ark Verma
Department of Humanities & Social Sciences
Indian Institute of Technology, Kanpur

Lecture – 04
Knowledge –III

Hello and welcome to the fourth lecture on the course, Introduction to Advanced Cognitive Processes. I am Dr. Ark Verma from IIT, Kanpur. In this week we have been talking about language in the first lecture on we have been talking about knowledge. In the first lecture, I talked about concepts and categories and I talked about the various approaches people have taken to categorize the world into these boxes and you know which helps or which warrants them a better understanding of the word.

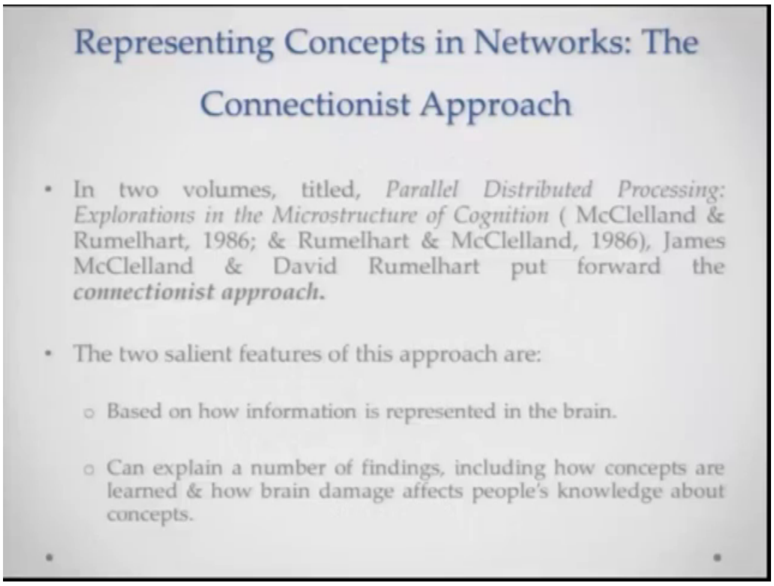
In the last lecture, we talked about particular theories which have tried to comment about how these concepts are organized in the human brain. We talked about the semantic networks theory proposed first by Collins and Quillian in 1969 and we saw that how these particular concepts are represented by nodes in these particulars in this particular semantic network theory. We also saw these nodes are linked to each other by virtue of these various links by virtue of properties they share you also saw that each of these particular nodes carried some information about the concepts they are talking about, say for example, if I might remind you we were talking about concepts like cannery, robin, sparrow, etcetera.

We were talking about the fact that canary at one level is a basic level concept, canary is connected to a slightly more general concept that of a bird and a bird at the bird note we have specified that a bird can fly, it has feathers and at canary level we have things like canary can sing, is yellow in colour etcetera. You also saw that there is a more general level that from bird you can go to animal from animal you can go to a living thing and all of those things are arranged very neatly, very nicely into a particular network which also represents how these concepts or which also let us say gives us a good analogy of how particular concepts may be a represented in our brain.

Obviously, there was no at the time this theory was proposed there was no real and neuro-scientific evidence to support or refute this theory, but again this was a good theory to begin with and to begin talking about organization of concepts in the human brain. We saw that this theory also had some you know some flaws to it and that those flaws were removed in a later theory which was given by Collins and Loftus which, however, although removed most of these doubts created by the earlier theory most of the problems that the earlier theory could not handle. But, in some sense that theory became too powerful to be refuted by you know empirical evidence as Johnson Laird pointed out.

Now, in today's lecture we will be talking about a different a slightly different approach to organizing concepts or organizing these categories in the human mind.

(Refer Slide Time: 03:09)



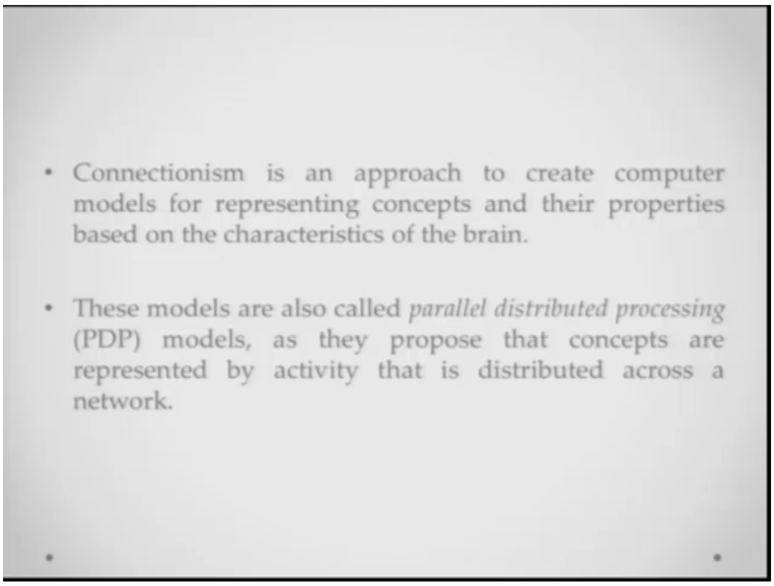
This theory is referred to as the connectionist approach to a representing concept. The majority of the major breakthrough or groundbreaking work in this particular regard was actually published in 1986, when the book parallel distributed processing explorations in the microstructure of cognition came out.

The authors were James McClelland and David Rumelhart and they actually in some sense began what is called the connectionist approach to cognition. They have worked a lot about cognition and various aspects, but in today's lecture we will be focusing more about how these are connectionist networks represent or store concepts and particular categories.

Now, there are 2 salient features of this approach. First is, based on how information is represented on the brain. This theory almost appropriates the way knowledge might actually be stored in the human brain. So, the typical neuroscientist explanation is that knowledge is not really stored at particular sites in the brain, but it might be distributed across the various neurons of the brain.

The connectionist approach or the parallel distributed processing approach which established by McClelland and Rumelheart takes this explanation and tries to computationally model it through what is called a connectionist network. So, it basically try to resemble as much as possible how particular concepts would have been arranged in the brain or are probably arranged in the brain. Also, this theory can explain a wide range of findings including how concepts are acquired how concepts are learned also findings about how knowledge about particular concepts might be damaged suppose, there is a brain injury or there is some other factors that are that are you know pitching in. So, again this is also one of the very popular, one of the very powerful theories a about organization of concepts that has been around for a long time now.

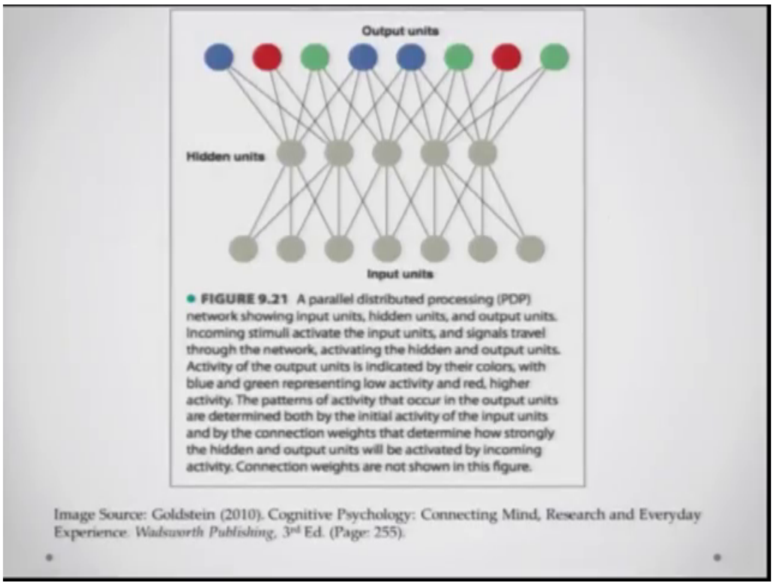
(Refer Slide Time: 05:08)



Now, a little bit about connectionism. Connectionism basically, is an approach to create computer models for representing concepts and their properties and it basically is based very closely to the characteristics of the human brain. So, in that sense you might say that this is more believable and the results that this the connectionist models would get might resemble in some sense as to a human processing, as to how humans might be processing these concepts.

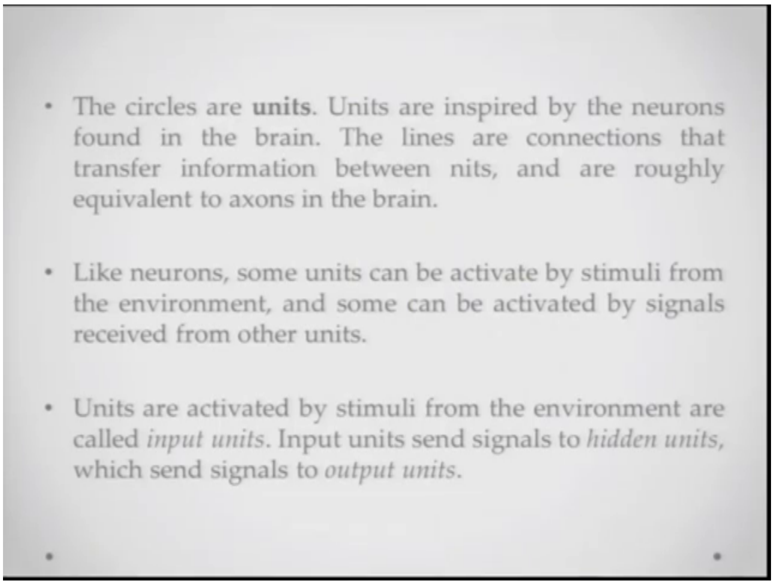
Also, these models, again as I already mentioned these models are called parallel distributed processing models, because they propose that concepts are represented by activity that is distributed across a network. Now, this is a major change of stance from the earlier theory we were talking about yesterday. We were talking about the semantics networks theory by Collins and Quillian, very particular concepts were represented by particular nodes and those nodes might be you know linked to each other. So, canary is basically stored in as one node. In the distributed processing approach or thus models which we talked about concepts might be represented only as a pattern of distributed activity across a wide number of cells or a wide number of neurons.

(Refer Slide Time: 06:19)



So, if you can see here this is one of the examples of these connectionist theories. You can see, then at the bottom there are input units and then there are hidden units in the middle and then there are output units at the top. So, the idea is that if you have to give some stimulation you probably give that at the input units some kind a of processing will happen at the hidden units while the output units would come with certain kind of output, a categorization task, a naming task or something very similar to that.

(Refer Slide Time: 06:48)

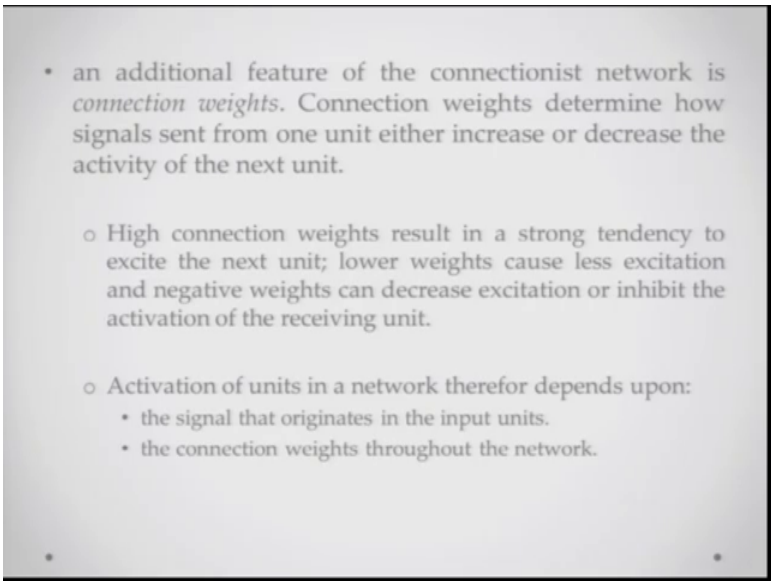


Now, how does this reverse these, a circles as you saw here I already told them. These are referred to as units and these units are basically inspired by neurons found in the brain. So, these lines which you see here these lines are actually connections that transfer information from a unit to a unit. Just like different neurons are connected to each other by axons. So, this is one; the second is, as in the case of neurons some units can be activated from the stimuli receive by the environment, some units can be activated by the signals you see from other units. Again referring back, if you see the input units have scope of being stimulated by the environment, you might present a question you might present a stimulus you may show similarly a picture a word anything this will cause activity in these input units.

Now, after this these input units will probably excite off the hidden unit. So, you see these input units are linked with lines to all of these hidden units and then the hidden units do some kind of processing and then they pass on this processing to the output units this is how this connectionist thing is working now units activated by stimuli as I have already again mentioned from the environment are called input units. Input units send their signals to hidden units where the processing is probably going to happen and these hidden units will send whatever that the outcome of their processing is going to be to the output units.

So, again just have a look at this and it is all very simple. There are input units, there are hidden units, and there are the output units. Input, where information current function goes from the input to the hidden and it goes from the hidden to the output unit.

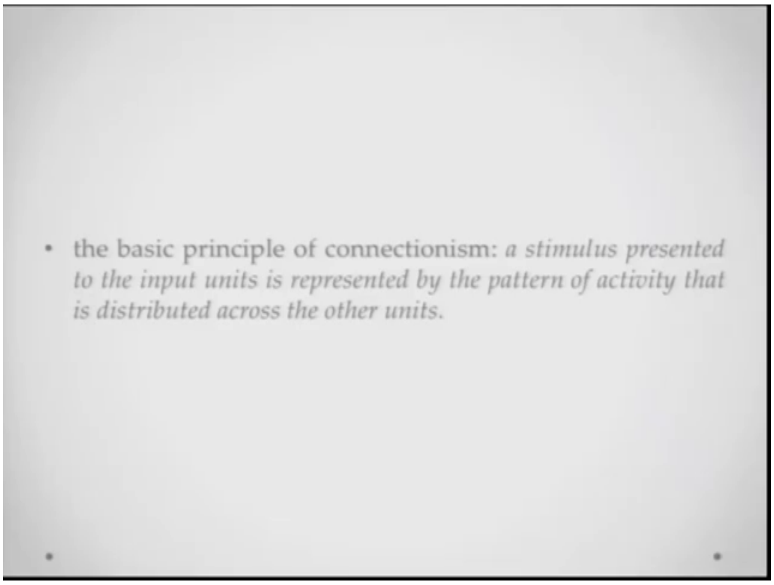
(Refer Slide Time: 08:32)



Now, an additional feature of the connectionist network is something called connection weights. So, how strongly or weakly each of these units might be connected to each other. Now, this connection weights basically determine how signals from one unit either increase or decrease the activity of the next unit. So, high connection weighs if two units are connected with heavy connection weights, they will result in strong tendency to excite the next unit. Lower connection weights basically, will cause less excitation and negative weights would basically lead to decrease in excitation or inhibition of whatever activation is there at the receiving unit.

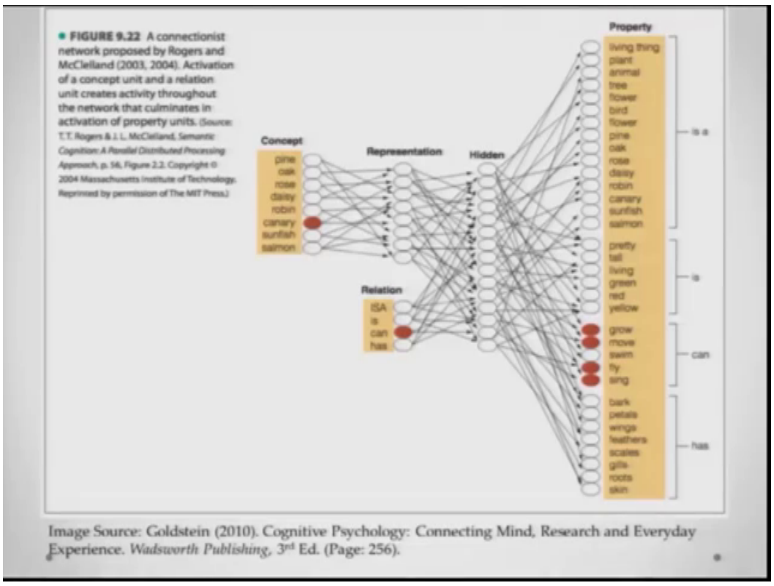
Now, activation of units now, in this kind of a connection is network basically then depends upon at least two things. First; the signal that is originating in the input units a signal a picture is shown and the picture has to be named. So, whatever aspects are representing that concept in that picture let us say a cat or a sparrow or whatever that will excite the input units related to processing that concept and then the other further excitation will be passed on to the hidden units. Depending upon the connection weights connection weights of these different input units information will be passed on or propagated in this network.

(Refer Slide Time: 09:53)



Now, the basic principle of connectionism if you got the idea by now is a stimulus is presented to the input unit. This is represented by a pattern of activity that is distributed across others. So, it is not like the one unit here will receive the concept of canary and the other unit will take on this activity as a whole. The idea is that whatever stimulus is presented to the input units it will be represented by a pattern of activity and this pattern of activity will be distributed across a variety of a large number of units.

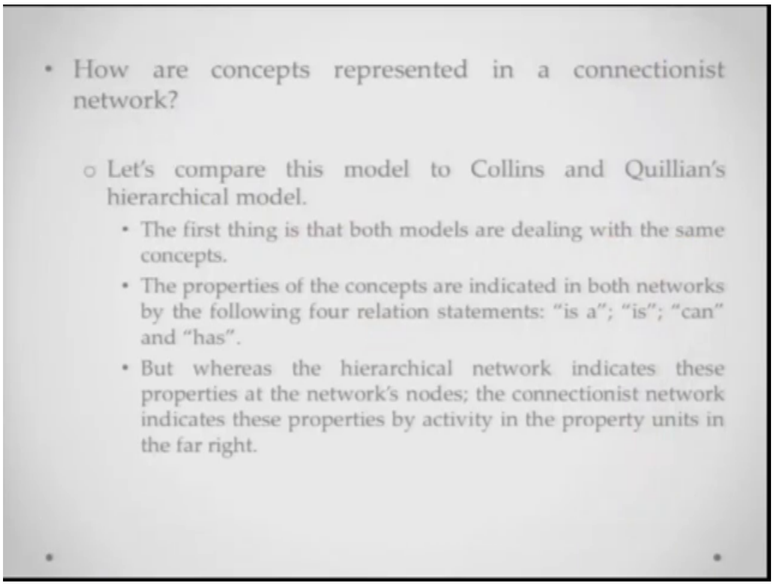
(Refer Slide Time: 10:32)



You can see here, this is an example of a connectionist network and you can see here that there are concepts, there are representations, again you can see the representation is rather distributed and then there are these bunches of hidden units and hidden units are actually connected to what are called property. So, there is where the processing is happening, the property units and properties could be variety of things living thing plant, animal, bird, flower, it is pretty, is tall, can grow, move, swim or fly or has barks, petals etcetera and you see properties are actually organized into is a links, is links, can links and has links. You can see these all of these kinds of links are actually represented in the relation units.

So, you can probably talk about let us say a pine is and then you can go to a tree or similarly you can start from canary can and it can grow fly move something like that you see the swim thing is not highlighted, because canary is a bird and obviously, cannot swim. Now, dead birds cannot.

(Refer Slide Time: 11:36)



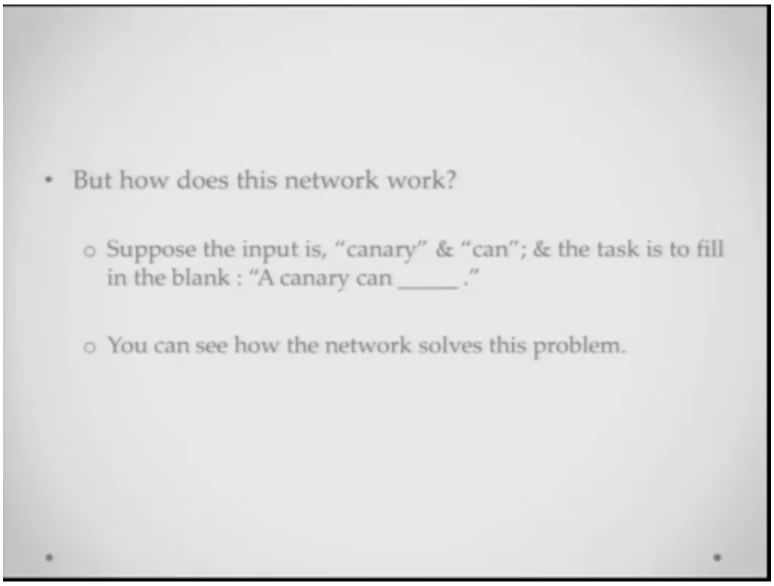
So, however, now, let us try and get a head around, how is this model going to work. How are these concepts actually represented in networks such as the one we were and now seeing. Now, let us compare this model before actually moving let us try and compare this model, in its basics to the Collins and Quillian's model.

The first thing you will notice and you might have already notice is that both models are actually dealing with very similar concepts of objects, plants, things etcetera. The properties of these concepts in are indicated in both the networks Collins network and this one by the 4 relationship statements and this is for relationship statement is can and has.

But whereas, in the hierarchical network which was Collins in Quillian's network these things are these properties are indicated at the networks nodes and the connectionist network indicates these properties by activity that is distributed in the property units at the end. So, you see is a, has etcetera in the Collins and Quillian's model was actually all denoted at a particular node. So, at the bird node as I was mentioning in the beginning of this lecture at the bird node you has feathers, can fly, can move about things like that, but here all of those properties are distributed at the end in the property unit.

So, there are many properties and there are many relation units at least four of them and there are many concepts. So, each concept via each of these relation links can be mapped onto this large array of properties that you see.

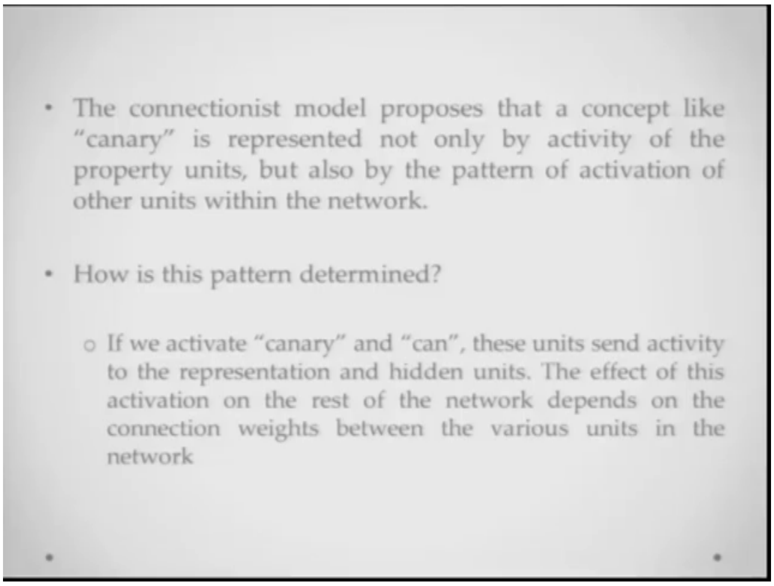
(Refer Slide Time: 13:25)



So, if you a take into account whatever I was just describing about this model, let us try and see how a particular example would work through this network. Suppose that the input is canary and the relation statement that we are beginning with is can. So, you want to basically you ask the network to fill in the blanks for this statement a canary can and then there is a dash now dash will basically it needs to be filled. So, you need to see how this network will solve this problem.

Again, this is the network you can see you start moving from canary, then you come to can and then you can see a from can it goes to let us say 4 out of 5 properties that are linked with this relation statement can. So, the 4 properties are it can grow, it can move, it can fly, it can sing, but it does not really light up swim because probably that is what the process and the hidden units has told us.

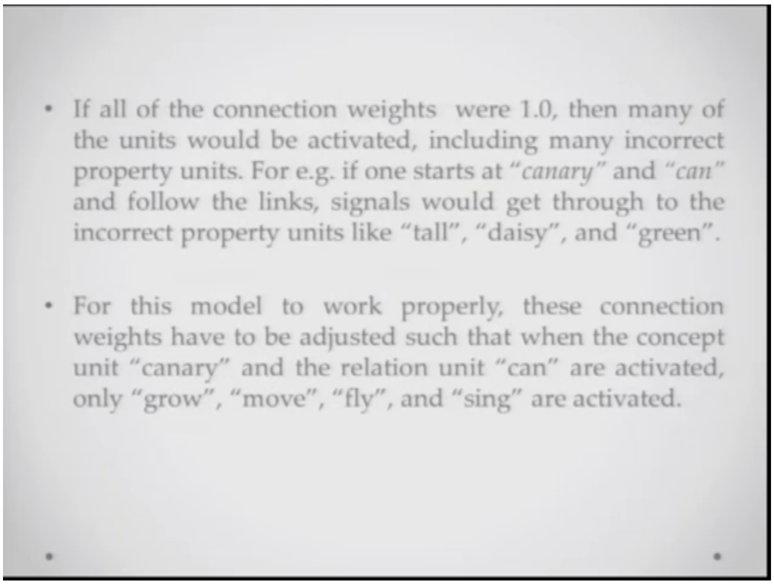
(Refer Slide Time: 14:19)



Now, the connectionist model it proposes that a concept like canary is represented not only by the activity of property units, but also as a pattern of activation of other units within the network. So, it is not only that there is only one concept in the entire network or and there are just properties associated with that concept, there are other concepts as well and sometimes these other concepts and their processing might be able to you know a influence the processing in this model.

Say, if we can participate if you want to determine this or if you can see, how is this pattern going to be determined. So, let us say if you can activate cannery and you can activate can and these units will send activity to the representation units and the hidden units the effect of this activation on the rest of the network will basically depends on the connection weights between canary and can and the various units in the network. Let us try and get slightly more into this concept of connection weights.

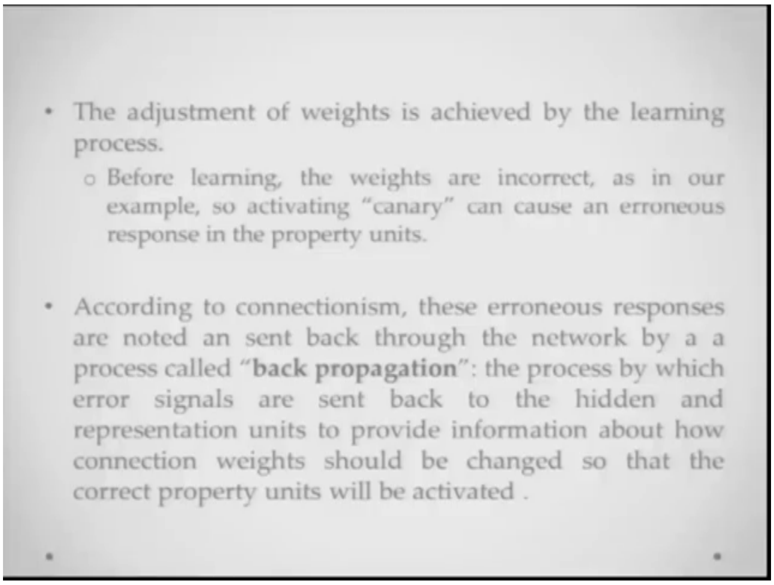
(Refer Slide Time: 15:13)



Let us say the connection weights of all of the concepts in the same network were 1 is basically 100 percent. Everything is connected to each other and all the information is going to go to all nodes then what will happen is then many of the units will get activated. Many of the units including many incorrect property and if one starts at canary and has the relationship statement can a it might follow all the links and it might also go to you know things like canary, can, tall, daisy, green again these are incorrect property units, that is not how the fill in the blanks will be done.

So, for this model to work properly these connection weights would needs to be adjusted. So, the connection weights basically determine how the output is going to be generated and. So, these connection weights will have to be adjusted in such a way that when the concept unit canary and the relation unit can are activated it only leads to these 4 units that you know represent what a canary can do grow, move, fly and sing. How does this connectionist network learn to come up with this final, what is going to happen, what kind of learning is taking place? Let us look at that a bit more closely. Now, that is I was talking about if the connection weights were 1 you know everything is going to get activated and lead to so many incorrect outputs.

(Refer Slide Time: 16:38)

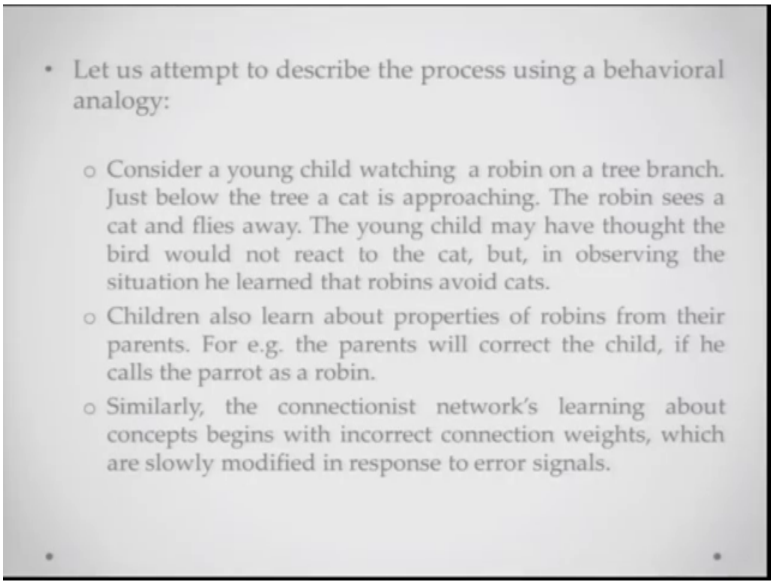


If you have to achieve the correct connection weights it has to happen by a process called learning before learning the weights are, obviously, incorrect as in our example we can we showed that canary and can cause act erroneous outputs like things like tall, daisy etcetera. According to connectionism, now, I am going to talk about how our learning is going to happen according to connectionism these erroneous responses like the ones we just saw, are noted and they are sent back through the network by a process called back propagation.

Now, what is back propagation? Back propagation basically is the process by which error signals suppose a this network has started learning and you are kind of repeating this process of getting an output and you try and decide that I will keep repeating on this process a till a correct output is there. So, every time an erroneous output is given these erroneous outputs are noted down and they are sent back to the hidden units they are sent back to the representation units with this information that this is wrong, this is wrong, this is wrong.

When this information reaches there, when it information reaches the hidden and the representation units the connection weights are adjusted. So, it is almost like you know you want to a yeah children learn a lot by correction. So, if a child kind of makes a particular mistake, you tell the child that this is not correct. Then next time it is more probable than not that the child will have corrected its mistake, it is pretty much happening in the same way in this network. So, the back propagation process notes down the errors senses it back to the hidden and senses back to the hidden and representation units and they are basically a the connection weights are now adjusted. So, that a correct property node will only be activated.

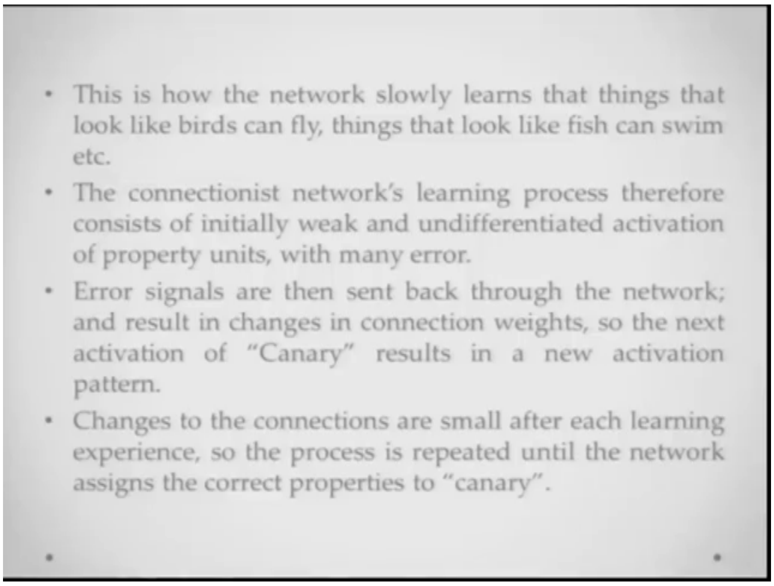
(Refer Slide Time: 18:31)



So, let us attempt to kind of again something I was saying, consider, this is an example from Goldstein. Child is watching a robin which is sitting on a tree is branch and then just around that time cat is coming and the child does not really know already that robins because there birds they avoid cats because they do not want to get eaten. Now, the child does not know this and the child is kind of waiting and it is kind of looking at the robin while the cat is approaching. As soon as the cat approaches the robin the robin flies away. Now, the child here whose observing this learns at least one thing he learns that look cats that robins do avoid cats because they do not want to be eaten.

Now, children do this a lot children learn about properties of robins from their parents as well. So, this one was observation, this one was something that the child learned in action, but their other is also that in which children learned something. They are often told by parents or teachers or by books. Similarly, the connectionist network also begins with a certain kind of weights; begin the certain incorrect connection weights. So, something that is already some residual connection weights are there, some activation is there. So, and these connection ways are slowly and eventually modified in response to whatever error signals are coming up.

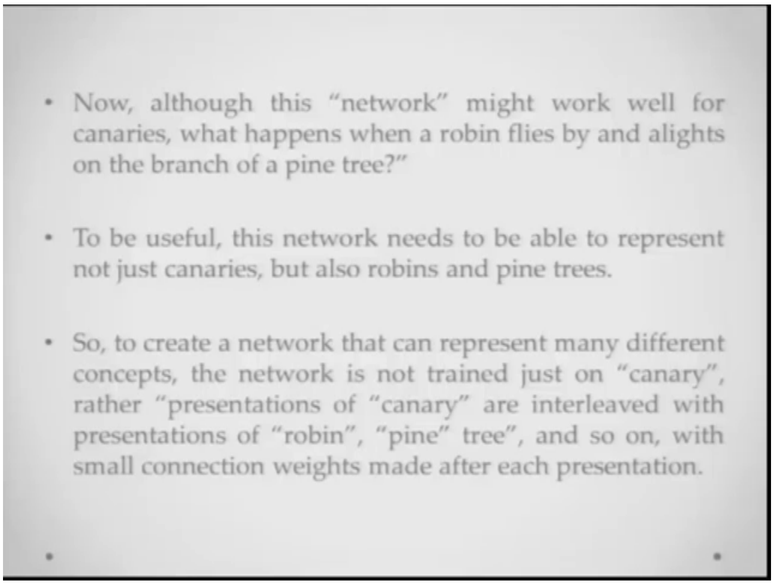
(Refer Slide Time: 19:57)



So, this is how trial by trial the network slowly learns that things that look like birds can fly and then things that look like fish can swim etcetera. Now, the connectionist networks learning process therefore, consists of initially weak and undifferentiated activation of property units everything is getting activated, you have no idea about you know when the network is going to give you correct input it will correct output, it will give. So, many errors, but as and when after every trial these error signals are being sent back through the network these error signals are then leading to changes in correction weights, changes or modification in the connection weights.

Every time the connection weights are adjusted so that the next trial is better, the next trial leads to less error and more correct response. So, activation of canary should lead to a different kind of an activation pattern, after a particular number of trials. Now, one of the properties of the connectionist network is that changes to these connections are slightly smaller, they are not very big. It is not having that after one trial automatically the second trial there will be almost correct output. So, these changes to these connection the changes through these connection weights that is are small after each learning experience. So, that the process is repeated again and again until the network assigns the correct properties to canary.

(Refer Slide Time: 21:25)

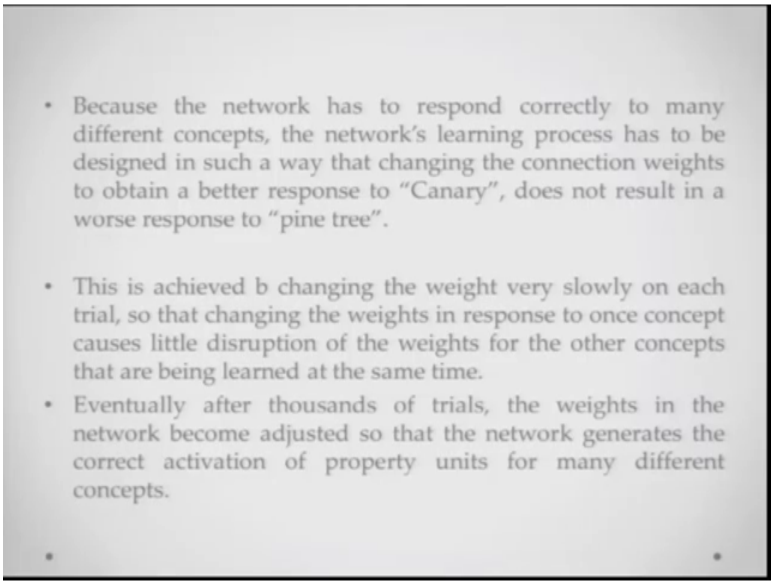


Now, although the network might not work well for canaries it might also happen that it does not really work well for other related concepts. Suppose, after a number of trials 5000 trials, 10000 trials the network can correctly identify canary and it is properties, but if robin comes and sits next to the canary, how is the network going to react to the robin.

Now, to be useful to be any useful and to be able in some analogous represent human conceptual organization this network would need to be able to represent not only canaries, but also robins and also other kinds of trees and animals etcetera. So, to create such a network and to create a network that can represent these many different kinds of concepts this network is not usually just trained on specific concepts like the canary or the robin or just the cat.

The presentations of canary, suppose you want to teach the concept of canary are also interleaved with other related and sometimes unrelated concept canary will be also presented with robin and a sparrow and parrot then maybe pine trees and other kinds of trees and as soon as the network is learning the network has to learn to classify each of these things correctly. So, small connection weight changes have to made after each presentation of so many of these concepts so that this network learns to correctly classify all of these concepts at once.

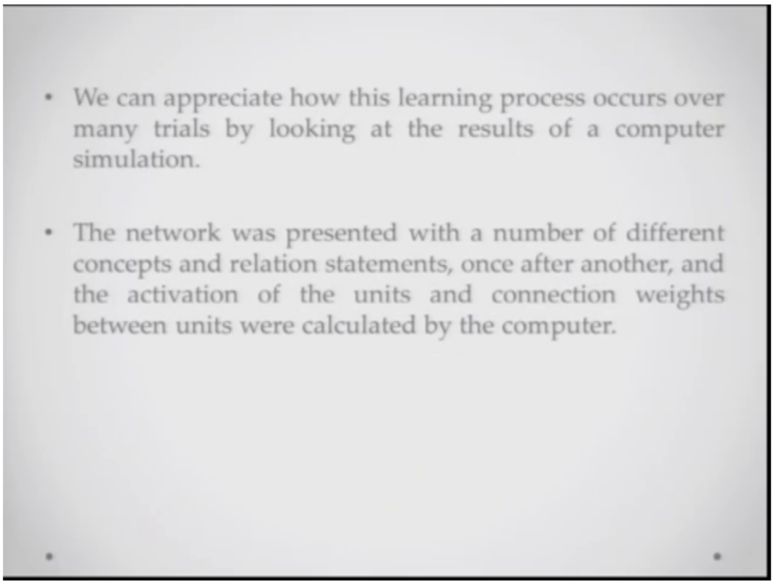
(Refer Slide Time: 02:57)



Now, because the network has to now respond correctly to so many different concepts the networks learning process has to be designed in such a way that changing the connection weights in response to let us say when it is learning about canary, it should not result in worse responses to let us say a pine tree or a robin. So, how do you ensure that the network is learning all of these concepts in a way and none of these learnings are interfering or making each other worse.

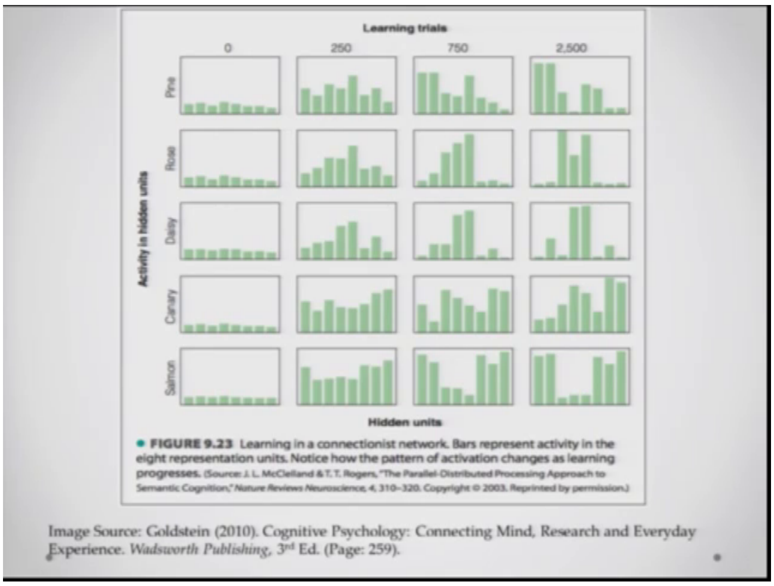
This is achieved by changing the weights very slowly. On each trial that connection weights are adjusted by a very small amount so that then when the connection when the responses to one concept are getting better it will cause very little disruption to responses towards other concepts that are again being learned at the same time. Eventually, with a lot of practice with thousands and thousands of trials the weights in the network become adjusted such that the network can generate correct activation of property units for so many different concepts at the same time.

(Refer Slide Time: 23:53)



We can appreciate actually we can have a look at this learning process and how this learning process occurs over many trials by looking at the results of a computer simulation then the example I am going to take is one where a network was presented with the number of different concepts and relations statement one after the other and the activation of the units and connection weights between these things were calculated by the computer. So, this is a typical example of the connections network.

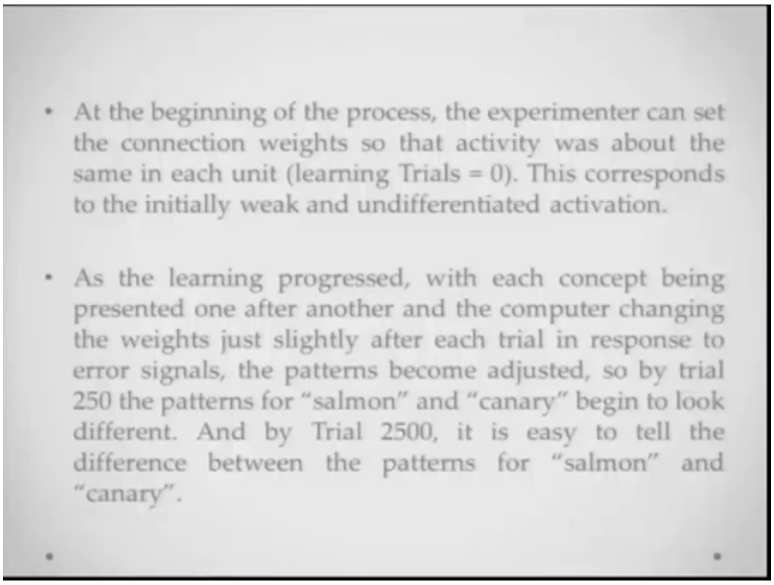
(Refer Slide Time: 24:24)



And, here in is basically is happening trial after trial. So, you see at trial 0 and then the number of trials is equal to 0, there is almost no difference between how the network is categorizing many different concepts like salmon and canary and daisy and rose and pine. But, you see as you go ahead in the number of trials and you go from 0 to 250 trials there is some change you go from 0 to 750 trials there is some more change, by the time you reach from 0 to 2500 trials, you will see that the count that the network has learned to classify all of these concepts separately.

So, you will see there is very little similarity between how the network is dealing with canary and how is this network dealing with salmon.

(Refer Slide Time: 25:16)

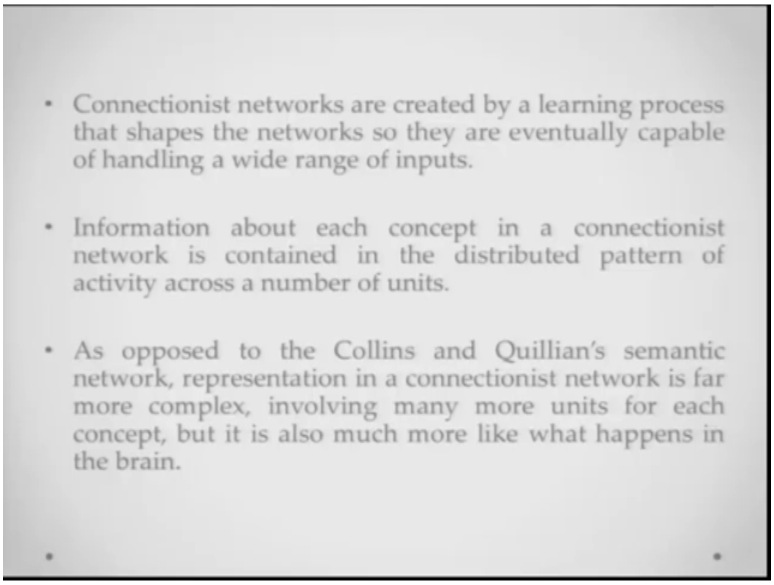


So, how is this happening? At the beginning of the process, the experimenter can you know set up some arbitrary can a connection weights so that the activity was about the same in each unit. Here, we are at learning classes of 0 this is basically what I was referring to earlier as weak and undifferentiated activation. So, anything that comes is kind of treated in pretty much the same way.

As the learning progresses with each concept being presented one after other again and again and the computer kind of changing weights very slightly after each trial in response to error signals by the time trial 2500 reaches the can network has learned here or you can see to classify these concepts very differently and probably come up with correct classifications.

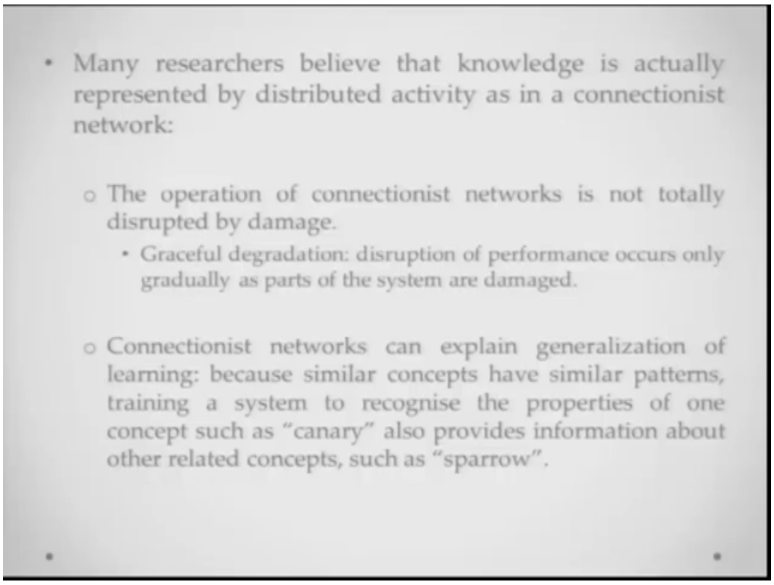
Now, these connections networks are created by learning are created by a learning process that shapes these networks so that they are eventually capable of handling a wide range of inputs. Information about each concept in a connectionless network is contained in a distributed pattern of activity across a number of units. So, you can talk about so many different birds and animals you can talk about so many different possible relations that are there and even talk about so many different properties has feather, has fur, has 4 limbs, has 2 ears, has whiskers you can think of you know a large area of property units. All of these are disputed across a number of finishes. So, it does not have one unit or one node for a particular property or particular category.

(Refer Slide Time: 26:48)



As opposed to the Collins and Quillian's semantic network, consolidation in this connectionist network is far more complex and it involves many more units for each concept, but it is also like much more advantageous and it is also a much more like the processing or the representation in the brain is happening. So, this is actually one of the major advantages of the connections network because as we know, and if you have referred to the earlier course we do not talk about memory for example, or perception for example, to be stored in a particular site in the brain. You are talking about memory as distributed across the very variety of areas in the brain. Similarly, we are talking about concepts being stored across a variety of areas in the brain.

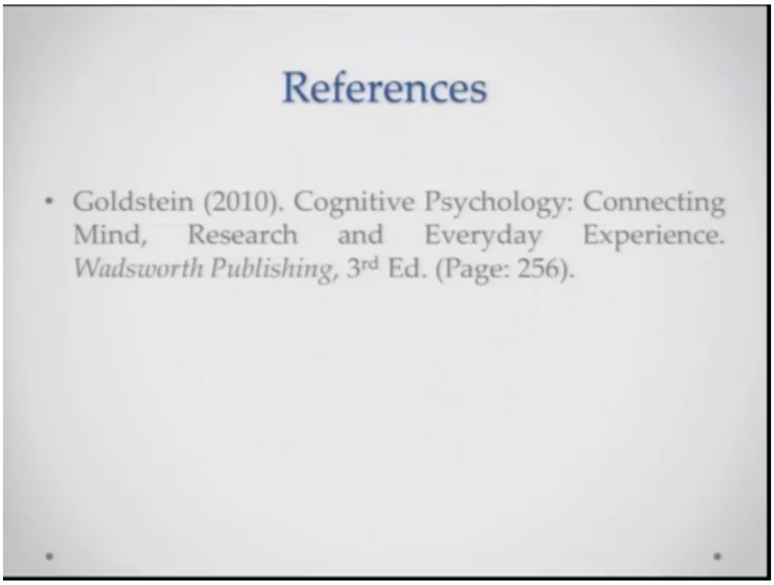
(Refer Slide Time: 27:31)



Many researchers actually believe that knowledge is in these connections network a really very closely reflects how the knowledge is stored in the human brain. One of the examples could be that the operation of connections network is never totally disrupted by damage. It basically gets disrupted very slowly this phenomena is referred to as graceful degradation. What does graceful degradation means? Graceful degradation means that if one aspect of this connection is network is damaged it at once does not break down or shut down the entire system. Very slowly gradually part by part, node by node the performance is degrading slowly and slowly. That is why, the reference to graceful degradation another aspect is that connections networks can actually explain the generalization of learning. You know in this world you do not learn about all specific animals, you do not learn about a dog and a cow and a cat separately you learn about all of them as animals.

So, because and again this is a very similar to what is happening in these connection issues because similar concepts have similar patterns. A training system which has basically learned to recognize the property as a of one concept such as a cannery or a robin or a sparrow also would be able to provide information about related concepts say for example, you train the system with canary it might be able to tell you little bit about the sparrow as well as the robin or as well as the parrot. So, again these two are examples in support of what the connections network might be talking about.

(Refer Slide Time: 29:04)



This is all from me about connectionist networks. We are still talking about knowledge in this week and I will talk about some other aspect of knowledge in the next lecture.

Thank you.