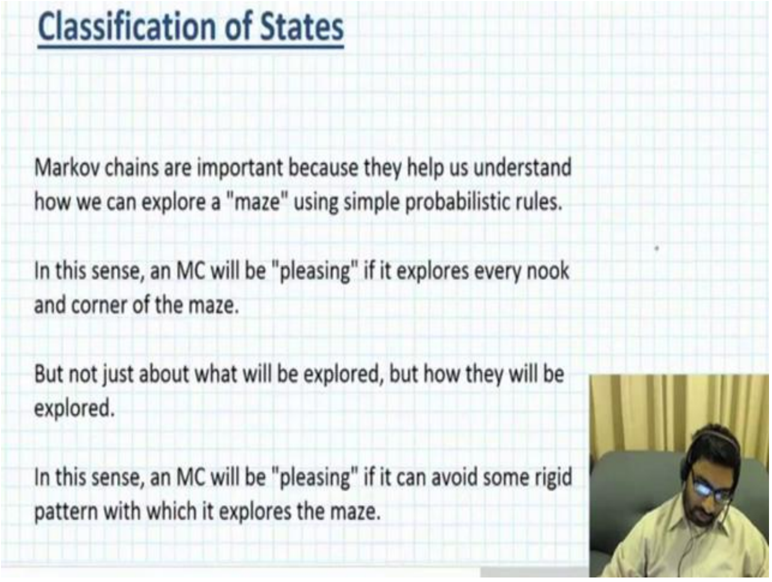
Algorithms for Big Data
Prof. John Ebenezer Augustine
Department of Computer Science and Engineering
Indian Institute of Technology, Madras

Lecture- 16
Classification of States

(Refer Slide Time: 00:16)



We have already seen what Markov chains are, and we have seen how they can be useful in the context of algorithms. And in particular, we saw how they were useful in analyzing very simple randomize algorithm for (Refer Time: 00:31) problem.

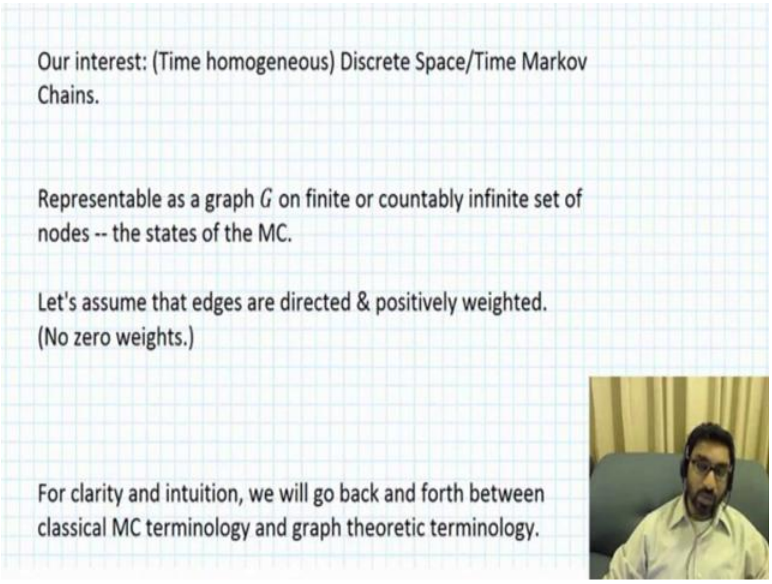
In general, Markov chains are interesting because they help us understand how to explore a maze using simple probabilistic rules, and that is what a appealing about them. And in this sense, a Markov chain is pleasing if it gives you way to explore the entire maze the word for it usually is the entire state space. And this it is a very natural way to explore it and it is somehow very appealing and turns out to be very, very useful. But it is not just about exploring because if it is a exploring the classic (Refer Time: 01:22), these do a perfectly fine job as well if you think of the maze is a graph.

So, there is something more to it. Markov chains have this ability to explore the graph in a variety of ways; it is not just about exploring, but how you explore that makes it a kind of exciting. So, in this sense, we are interested in Markov chains at that explore the maze, but not in some rigid fashion, but in some sort of a natural way and kind of not falling into parents, but exploring the state space in a variety of ways, so that is what is very appealing about Markov chains. And the rules that drive them are simple probabilistic transition rules; so it is very appealing.

And in order with this sort of a philosophy, behind which these Markov chains are designed, let us try to understand them a little bit more carefully. And for that purpose we are going to understand how the states are classified; and when you classify the states, we are also going to classify Markov chains and we are going to bring out some properties that make a Markov chains more interesting than others.

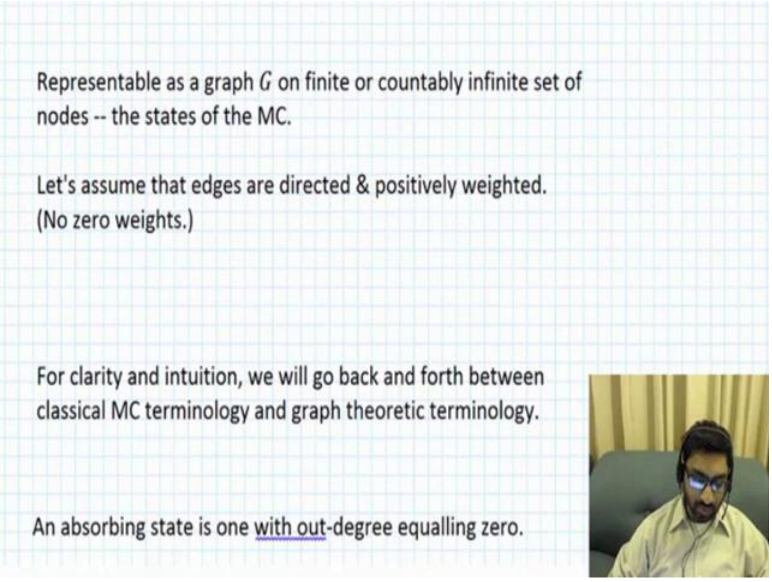
At this point, in our exploration, we are kind of thinking of the Markov chains is being interesting appealing and so on and so forth. Eventually, we will also get to a point where these Markov chains are not just interesting from a hypothetical standpoint, but also useful for in some concrete context. One thing to bear with is that we are going to go through a series of definitions, but keep this philosophy of Markov chains in the back of your mind, so that these are not just a sequence of definitions. But there is some method to this (Refer Time: 03:28) some underlying perspective that guides us in setting up these definitions the way they are.

(Refer Slide Time: 03:41)



Just to remind ourselves, we are interested in time homogeneous, discrete space discrete time Markov chains. During this lecture, when we talk about a Markov chain, this is the type of Markov chain that we are referring to and sometimes just shorten it to MC.

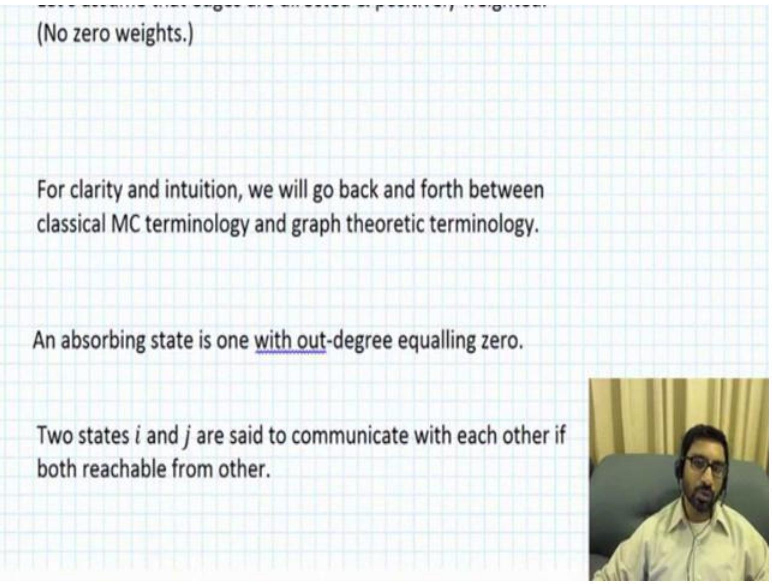
(Refer Slide Time: 04:02)



And also to remind ourselves a Markov chain can be represented as a graph in which the nodes or the states of the Markov chain and the edges are directed and positively weighted and those positive weights can be thought of as probabilities. So, if you add up the weights of the edges going out of a state, it has to be one.

And just to be clear sometimes you may not have any outgoing edge explicitly mentioned in the those states are called absorbing states; and essentially once a Markov chain reaches such a state, it is never going to leave that state. So, one way to think of it is these are states that have a self loop where the out edge the (Refer Time: 05:03) can be thought of as the out edge with probability one.

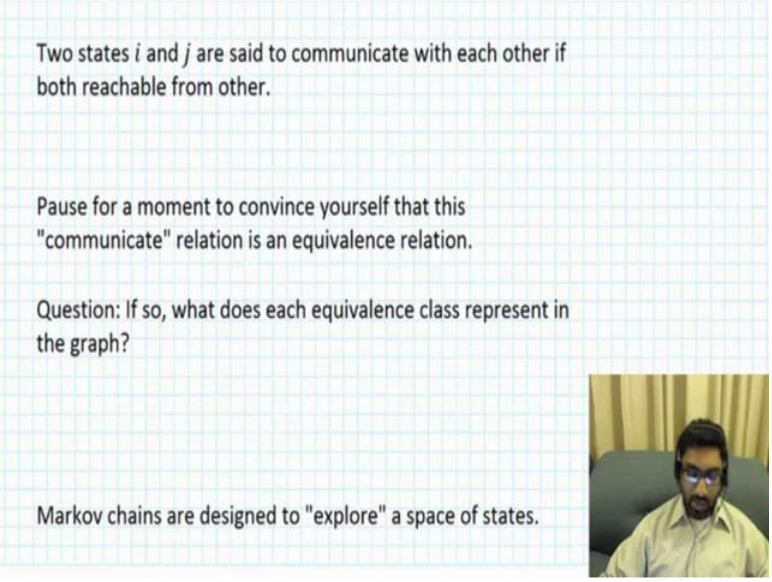
(Refer Slide Time: 05:13)



And throughout this discussion, we are going to go back and forth between the classic Markov chain terminology and the graph theoretic terminology. And for many of us including myself the graph theoretic terminology comes quite a bit more naturally than the Markov chain terminology, but they are fairly twine with each other.

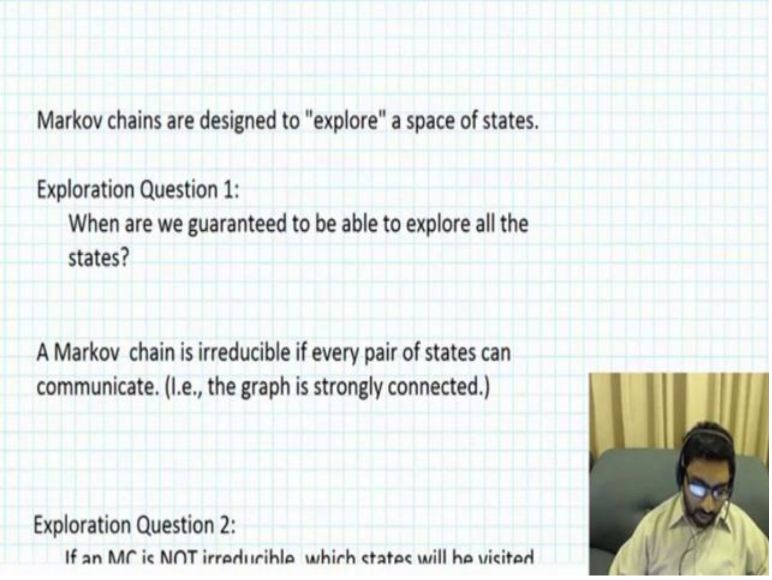
And so the point of view of gaining a better intuition about what is going on it is better to keep the better to relate Markov chains and graph theory the underlying graph theory perspective and keep going back and forth so that it reinforces our intuition. Two states i and j are said to communicate with each other, if they are both reachable from the other. This is a Markov chain definition, but there is a very easy and clear way to understand this from the perspective of a graph theory.

(Refer Slide Time: 05:55)



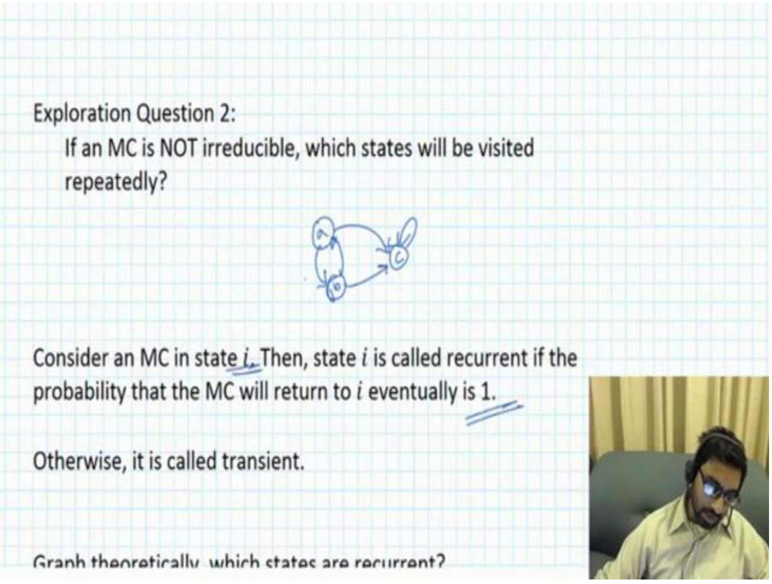
Remember the graph that represents a Markov chain well those two states that can communicate with each other essentially two nodes such that there is a path from a node say i to node j and also there is another path from node j to node i. So, this is what we mean by to communicating states of Markov chain, essentially they are connected at the in the underlying graph representation. With our goal of understanding Markov chains as tools to explore a space of states.

(Refer Slide Time: 06:53)



Let us ask a question. When are we guaranteed to be able to explore all the states? Since, the graph is directed; it is easy to construct Markov chains such that some states simply do not get explored and depending on where you start and so on and so forth. So, those Markov chains in some sense are well at least less interesting because you kind of miss part the graph. So, this leads us to the definition of an irreducible Markov chain and as one in which every pair of states can communicate with each other, and this corresponds to a graph that is strongly connected. So, from any state, you can reach another state and vice versa. So, irreducible Markov chain simply a Markov chain whose underlying graph representation is strongly connected.

(Refer Slide Time: 08:03)

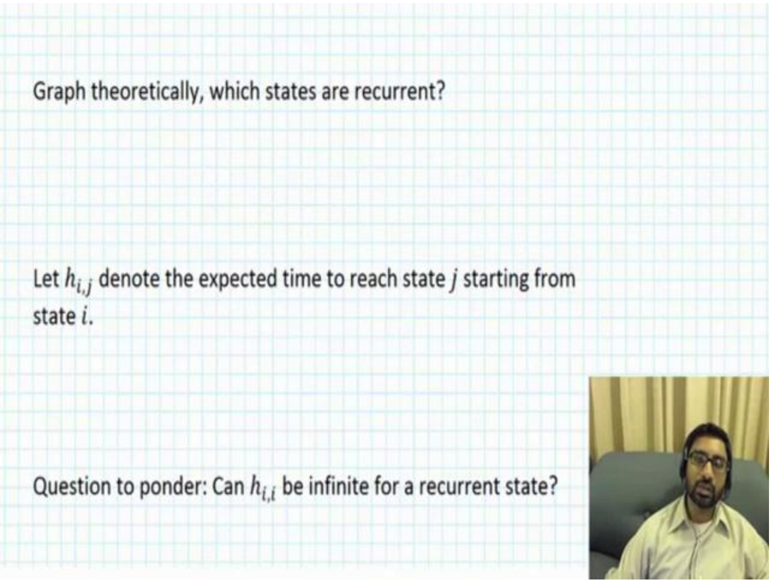


Let us ask our second question. Just to give some importance to Markov chains are a not irreducible. If an MC is not irreducible, which states will be visited repeatedly; in other words, are there some stage that will be visited once and may be once or twice and then you somehow stop visiting them is that a possibility. Well, of course, yes, if you have a directed graph, so let us make ourselves an example here, you will have to make up the transition probabilities, but well, let us see something like this. If you start at either state a or state b, you might be going back and forth a little bit, but once you go to state C, well C is an absorbing state, and there is no hope of coming back to A or B.

So, let us explore this ideal bit further. Let us consider a Markov chain in some state i, we will call that state i recurrence state if the probability that the Markov chain will return to i eventually is 1. So, in this simple example, once you go from A to C then there is no hope of returning back to A. So, if you are at state A, there is a certain probability with which you will never return back to A. So, then A is not a recurrence state; however, C is a recurrence state, because if you are at c, eventually you are going to come back to C.

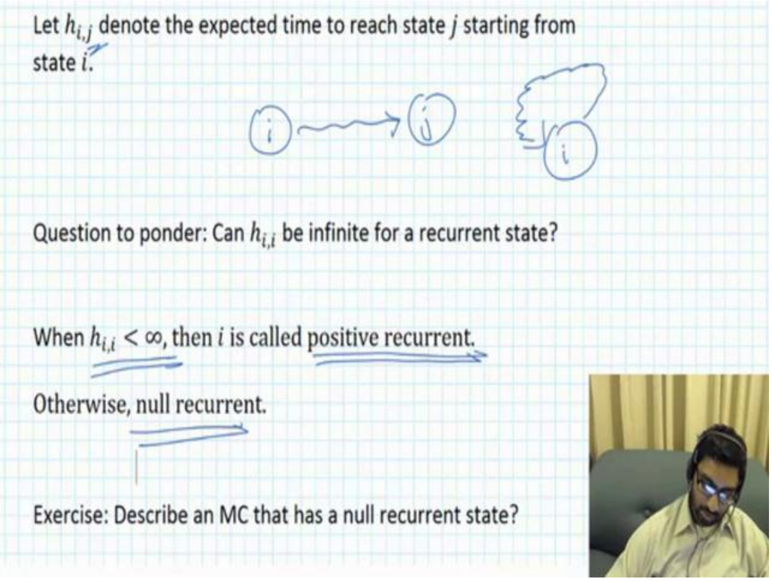
In this case, in fact, you will be at C at every time step, but you can design more interesting Markov chains where you can come up with states, so that you keep on returning back to that those are called recurrence states. Otherwise, they are called transient states because well the Markov chain may go through it, but it may not it may it may go into another part of the Markov chain from where it will never return back to (Refer Time: 10:25) such states are called transient. So, in this case, A and B are called transient states.

(Refer Slide Time: 10:33)



So, from a graph theory point of view, the question is what states are recurrent, and what states are transient. Think about this a little bit carefully, you may want to pause for little while, and answer this question to yourself (Refer Time: 10:48) because this will ensure that you, your intuition is going in the right direction. So, the questions clear if you are given a graph theoretic representation of a Markov chain, how do you figure out which states are recurrent and which states are transient.

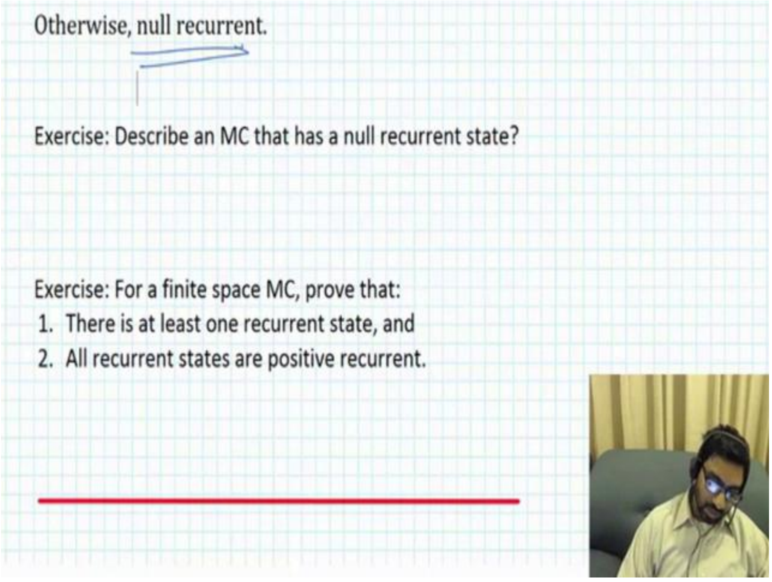
(Refer Slide Time: 11:01)



Let us use the notation h i comma j to denote the expected time to reach a state j starting at state i. So, the question to ponder is suppose you are at state i, how long does it take for you to get back to i. Remember we are talking about recurrent states, how long does it take for you to get back, and what is it on expectation, so that is this h i i. Can this h i i be less than infinite; of course that is quite possible; in such cases those recurrent states are called positive recurrent.

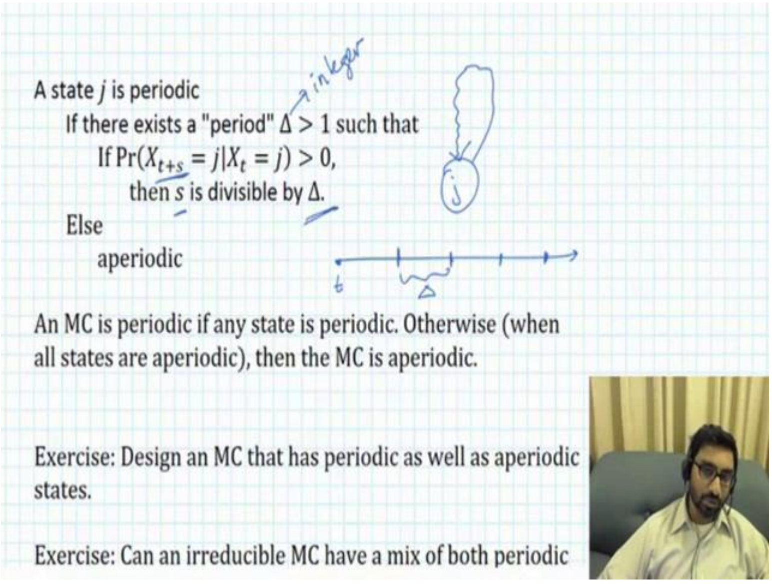
Interestingly, there are also examples where a state is recurrent, but it is not positively recurrent meaning the h i i is actually infinite. So, if you want to pause and think about when such a situation like that arise, and such states are called null recurrent. So, when h i i is strictly less than infinity then the state i is called a positive recurrent. So, your task is to describe of a Markov chain that has a null recurrent state. So, this might require you to think about an a little bit.

(Refer Slide Time: 12:57)



One more exercise, consider a finite Markov chain, now prove that this finite Markov chain has at least one recurrent state, prove that all the recurrent states are actually positive recurrent. In other words, a finite state of Markov chain cannot have null recurrent states. So, this actually works this exercise works as a goal to answering the previous question.

(Refer Slide Time: 13:30)

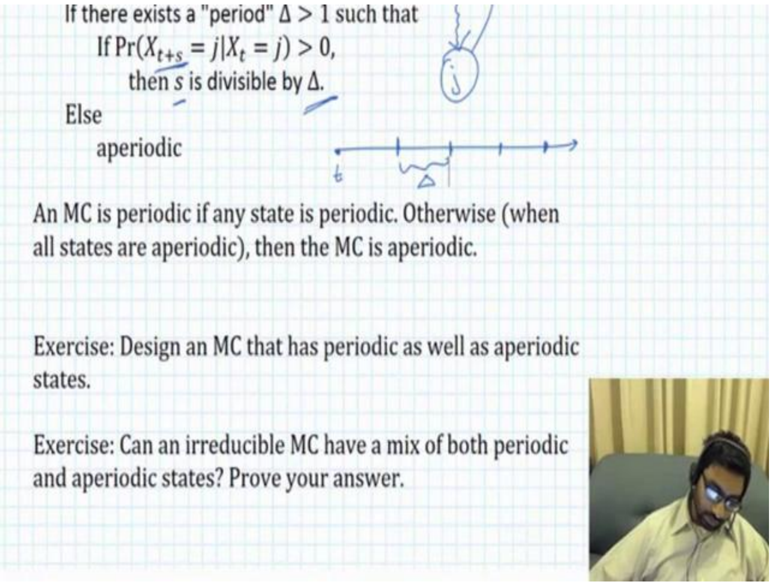


Recall that we are not just interested in how many what states can be explored and what states are not explored, we are also interested in ensuring that the Markov chain has interesting ways of exploring the state space. And in this sense, we want to avoid some rigid patterns and we are interested in Markov chains that have a nice fluid way of exploring all the states. So, from this point of view, we are going to study a notion called periodicity. A state j is said to be periodic, if there exist a period delta which is a an integer strictly larger than 1, such that if the probability given that you were at j at time t that you would be at j again at time t plus s is greater than 0 then this s is divisible by delta.

So, let us see if we can get a picture. So, we are talking about a state j, and we are at state j at some time t. And we are considering all future possible revisits of j, and those are at time say t plus s. And the only way you can have this probability of returning to j at some time t plus s is greater than 0, is that this s is divisible by delta.

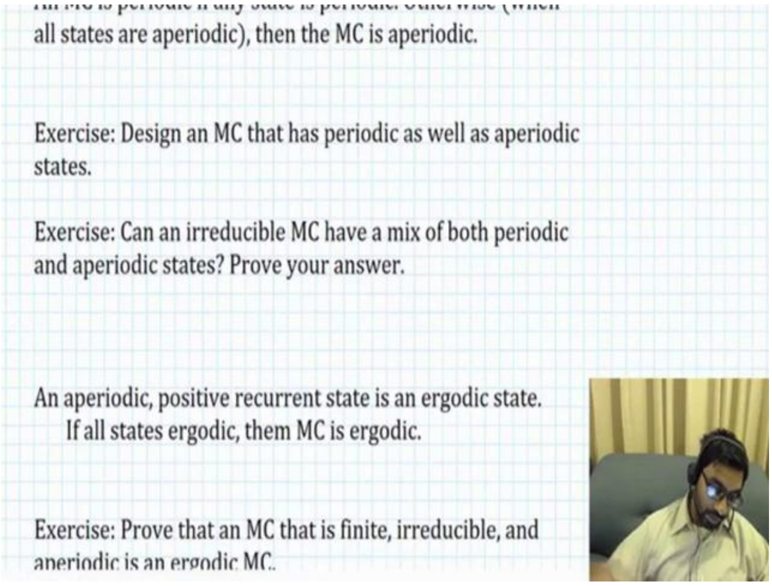
In other words, if you think of the timeline starting from t, the only possible points in time that you are likely to visit the state j again are all multiples of delta. And this kind of if you pause and make sure you understand the definition you will realize that there is some rigid pattern by which you are revisiting the state i state j. So, then the state j would be called a periodic state. If such rigid patterns are avoided then the state j is called a periodic. And a Markov chain is periodic if any state is periodic. So, the moment you have one state as periodic then the whole Markov chain is said to be periodic; otherwise, we call the Markov chain, aperiodic Markov chain.

(Refer Slide Time: 16:46)



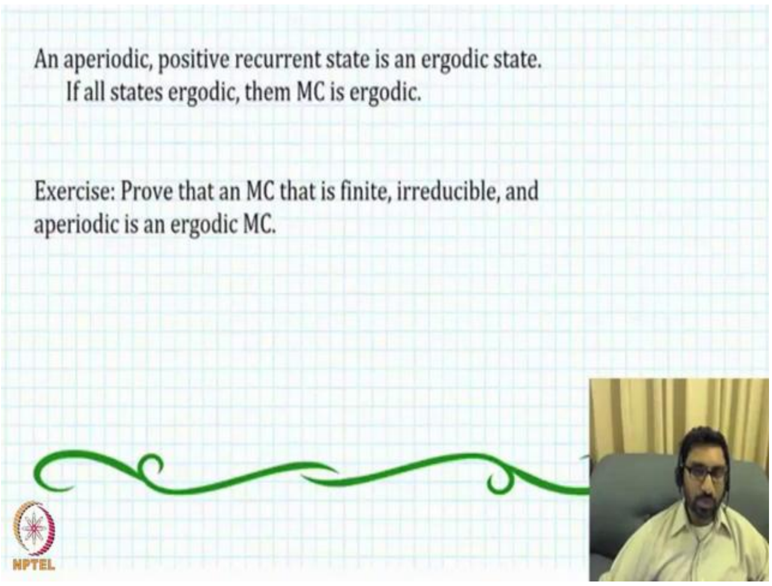
So, let us look couple of exercises. First, design a Markov chain that has periodic as well as aperiodic state. So, think about it carefully and come up with a Markov chain basically a directed graph with appropriate weights, such that it some states are periodic, while others are aperiodic.

(Refer Slide Time: 17:10)



And another question, consider an irreducible Markov chain, basically this is a strongly connected graph, can such an irreducible Markov chain have a mix of both periodic and aperiodic states? So, think about it, and make a claim, and prove your answer.

(Refer Slide Time: 17:31)



So, remember we are interested in recurrent states because well the Markov chain is likely to keep visiting them over and over again, but we do not want them to be visited in some rigid fashion and such states that are aperiodic, but positive recurrent are called ergodic states. And if all states in a Markov chain are ergodic, then the Markov chain itself is called ergodic. So, here is an exercise for you, prove there are Markov chain that is finite, irreducible and aperiodic is an ergodic Markov chain. If you think about it, you have actually already proved it in a previous exercise, so that will be the clue.

And with that, we end the segment on the classification of states in a Markov chain.