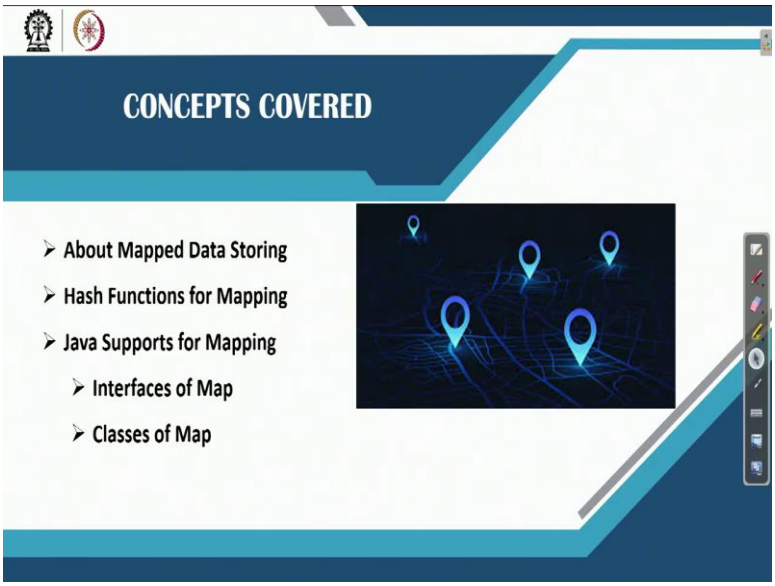
**Data Structure and Algorithm Using Java**
**Professor Debasis Samanta**
**Department of Computer Science and Engineering**
**Indian Institute of Technology, Kharagpur**
**Lecture 37**
**Map framework in Java**

We have studied most extensively on framework it is called Java collection framework parallel to this java collection framework, there is one very important framework also known, it is called the Map framework. Map framework basically is to serve on data structure it is called a table. Now, table is a very common one form of data and particularly in most of the IT applications, the table is there so, a collection of tables it is basically from what is called a database.

(Refer Slide Time: 01:18)

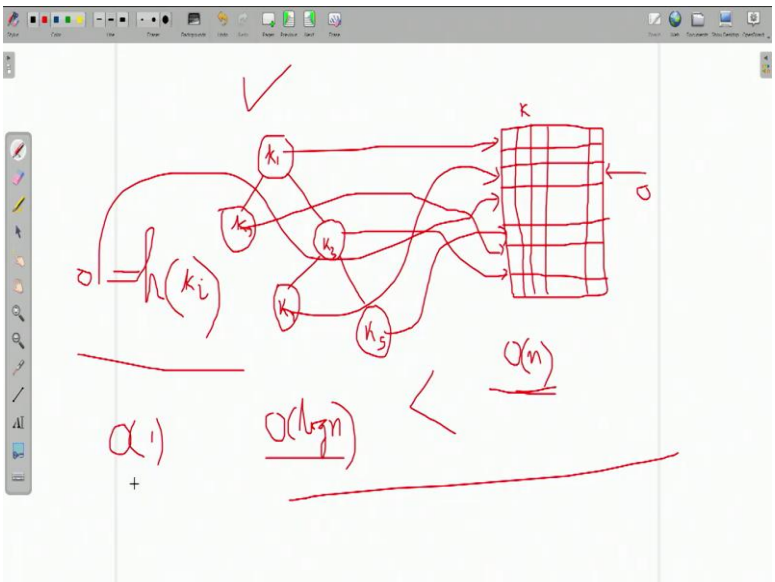



So, today our target is to cover this is very common, but extensively used data structure it is called a table and the map framework in java which basically supports all activities those are related to table structure. So, basically we will try to discuss about, how the java concept there is a map concept actually, that is used to store tables and manipulate tables and behind this manipulation, one important concepts is known, it is called a hash function so, we will try to learn about hashing or hash function and then, the map framework utility though, which is defined in java dot util package will be discussed in details. So, this is basically our plan.

(Refer Slide Time: 02:14)


Now, let us see what is the problem or other issues so, for storing data is concerned and how these issues can be addressed with the concept of what we are going to learn today, it is called a mapping.

(Refer Slide Time: 02:32)




So, I will try to give one idea about, this concept of the storing data now, consider there is a table so, a table is basically is a collection of records. So, we can say there is a number of rows so, each rows is basically is a record for example, there are say large number of students whose data needs to be stored in a database so, we can say a table can be planned for that for each record corresponding to a student's data so, this record is also in our concept of object

oriented programming, it is called objects so, a student object we can say. Now, as you know, each record is defined by a number of attributes so, these basically constitute is called the columns. So, there are maybe any number of columns are there so, these columns basically in the context of object it is basically called a field.

So, an instance of a student's class which basically constitute a record and they are each a field which can be considered as a unique there are maybe more than one fields also by which each record can be uniquely identified so, such a field it is called a primary field or we can say primary key.

So, let this is the field K is the primary field and these are that many objects we can say so, this is basically the third object so, it is an object O so, a table is nothing but a collection of objects and each object is identified by a field or it is called a key that is called a primary key now, so this is basically storing a table in a database, which is the common concept now, what is the problem, let us see.

Suppose, we want to modify a record for a given student so, what we have to do, we have to reach to that row where that student's record is available so, how we can reach so, we have to search the table and then, if records then, that record can be read and then, whatever the modification is required we can do it so, searching is a very common what is called operations, which is required in case of handling table.

Now, you see the searching can be done in many different ways. So, a trivial searching procedure is basically starting from the first record and compare the key of that record which basically our target record and go on until we find the record or we can reach at the end of the table but, this process you know, it is easy simple although but, it is not an efficient particularly when the large number of records are stored in the table so, what is the procedure?

We need one some mechanism by which this searching time can be minimised because, in a table there are a huge number of records accessing is require so, each time searching takes enough time then, definitely it will degrade the performance of the system so, one what is called the indexing mechanism is required.

Now, towards this direction I mean indexing, we can apply binary search tree which we have already studied. Binary search tree is basically an important data structure for the purpose,

what binary search tree does for us? So, for the indexing is concerned it is very simple you know, in case of binary search tree, there is a number of nodes number of nodes and all the nodes are arranged in a particular order is basically satisfying the binary search tree property.

Now, here each node basically store the key value off a record so, K1 K2 key value of record so, K3 whatever it is you can see K4 and so on so, the number of nodes is basically in the binary search tree is same as the number of records in this table now, K1 is basically you know, if we want to search a particular record having a particular key that has been the primary key then, we can take the confidence of these binary search tree to go fast.

Now, for this for each key it also store one additional information the information regarding actually which is the memory location where that particular record which key value is store in the node so, this means that each node store not only the key values but, in addition to a pointer, pointer to which record it is stored there so, it basically store a pointer so, suppose this is a pointer, which is basically pointing up a record which key value is key five and it is like this and so on so, this pointer is there and like this.

So, this binary search tree not only stored the key values but, also stored the pointers. So, suppose you want to access a particular record whose key value is K3 we can just take where is the K3 is there go, come here and then go to the record so, we can go it now, what are the advantages? Advantage is that the earlier searching which was simple and trivial it took order of n time that mean, n number of comparison that is required in order to find the particular record on the other hand, if we take the binary search tree then, it will take order of log n and since this order of log n is less than order of n so, we can say that, this procedure is far far faster than, this procedure that is that trivial linear searching.

Now, here basically you can see we have followed certain concept of indexing or there is a mapping there is an so, this basically the mapping that we have discussed but, there is another mapping this mapping is far far better, than both order of log n or order of n what is the mapping?

The mapping is that there is a function let this function be h these functions is called hash function and if the target key say key that is, we want to find a record having the K so, let K either i th is record we want to find now, this h function is planned in such a way that if we give the Ki value then, this hKi will give the address where, this record having the key Ki is

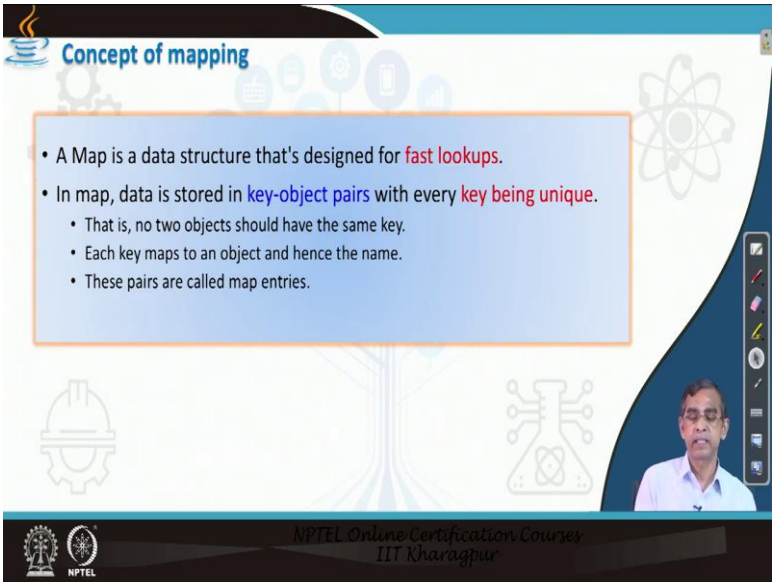
store so, these basically still that where this O is store so, using this will be able to directly go to this one.

Now, this is the one called hashing technique by which we can do it far far better because, this takes order of constant time it is independent of in whatever be the size of the tables, it is absolutely independent of that and therefore, this is far greater than, this one so, the concept is here is called a one hashing concept that is we can use it.

(Refer Slide Time: 11:03)

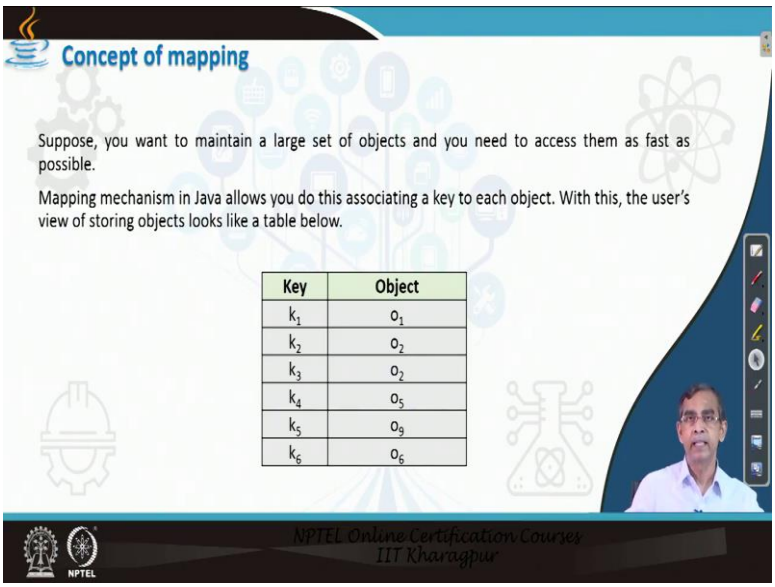

So, now today we will discuss about this concept of hashing in our study and then in order to manipulate the table so, storing data with mapping is basically using the hashing concept that we are going to learn.

(Refer Slide Time: 11:18)



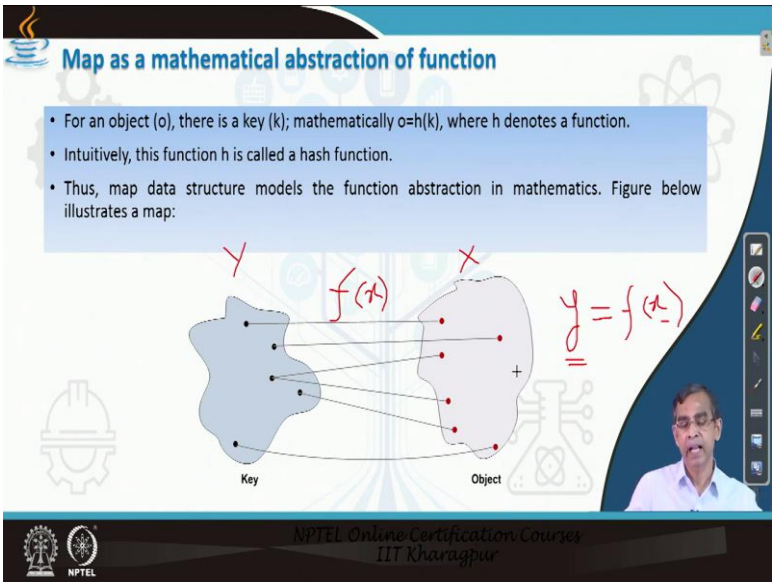

So, let us start about the concept of what is called a hashing concept it is there now, as we see here so, in case of hashing technique, there is a two things that do be in are involved on is that for each object and there is an association with a key value so, key object pairs are there and then key obviously should be unique that means there should not be two objects for which the key values are same, that means two different object is are there then, they should have the different key values are there so, it is basically very good for fast searching and each key is basically maps an object and this mapping can be often by a mapping formula, it is called a hashing concept is there.

(Refer Slide Time: 12:10)



So, that is what the concept is that for every object Oi there is a key Ky and they are for each object is mapped by means of Ky, Ki now, that is why this mapping is possible, this mapping is possible by means by a h function it is called a hash function.

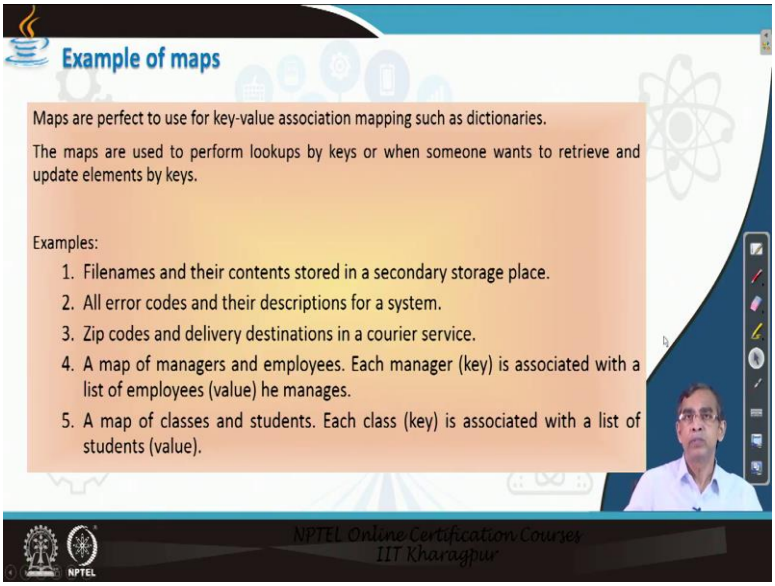
(Refer Slide Time: 12:28)




Now, this mapping if you see it is very similar to the function for example, y equals to a fx function so, what is that the theme, theme means that our x is basically dependent x is basically x is a independent variable and y is the dependent variable so, for a given value of x, there is a value of y it is not true that for a given value of x there are two values of y it is not the property of function but, for two different values of x they are maybe same values of y it is possible in case of mapping anyway so, this is a usual concept of mapping that is known in mathematics.

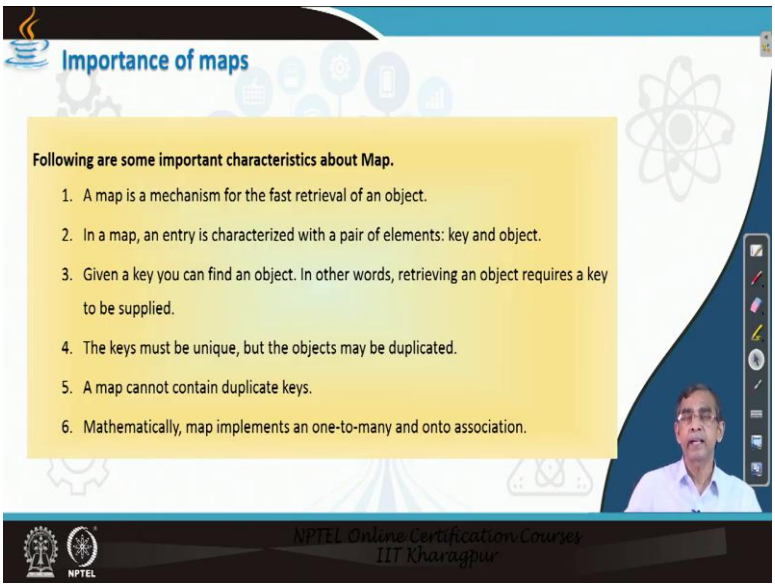
So, here mapping means if this is the set x and this is the set y is basically and this mapping is the function effects means how the x can be related to y and (())(13:31) it is there so, this mapping is basically the concept behind a hashing technique that we are going to use for our database or table manipulation.

(Refer Slide Time: 13:44)




Now, so we have learned about the mapping, mapping has extensively used many application in our usual concept of file structure that is used in computer memory, there is every file name has association to a file object actually which is stored in memory so, there is a mapping from the file name to file which is stored in hard disk or secondary storage and like this the mapping is every error for every error code there is a corresponding error mechanism routine and a for every country or every location there is a PIN code so, there is a mapping from location purchase PIN code and like, like so, there are many examples of mapping which we basically are familiar to.

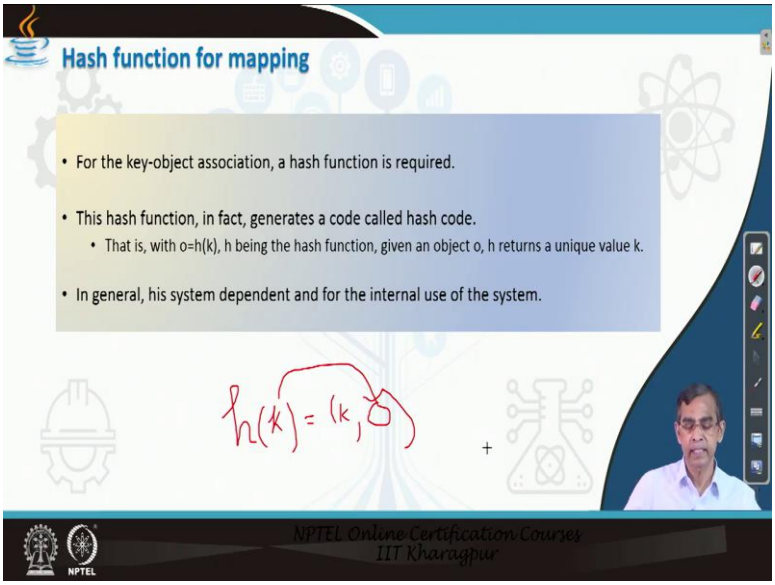
(Refer Slide Time: 14:30)



Now, this concept of mapping is very useful to help us in indexing a table so, is basically mapping is useful for fast retrieval of any object or object where basically represents a record and entry is characterised with pair of elements in our mapping mechanism as you have already mentioned about the record itself and the key value and mapping is basically by virtue of this mapping mechanism or hashing technique is there and as I told you the key must be unique as you know, in case of binary search tree and that is how we have assumed that there is a field for every record for which each record have that distinguishable or discriminated by that key value or field.

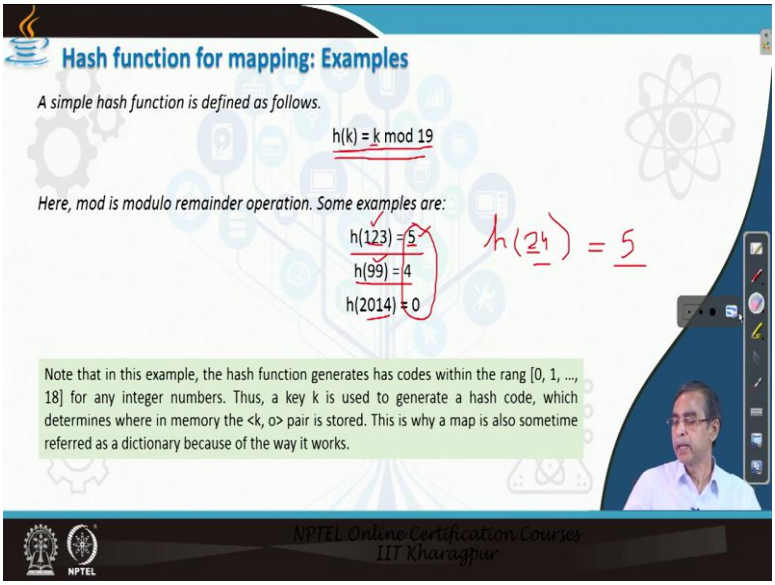
So, that is why it is unique and so, duplicate key it is possible but, we should not include that duplicate key is there so, a map should not contain any duplicate keys in that sense so, mathematically the mapping which we are going to use is basically on to on and on to mapping it is called the association.

(Refer Slide Time: 13:44)




Now, let us come to the concept of hashing as a mapping function now, so it is basically as I told you, in case of hash function so, there is any for any object there are two pairs key k and O so, these basically if hash function is there so, hk then, it tell that which is the object it is basically so, this concept is there so, there is a hashing technique is for this purpose only now, let us the come to some example so, that we can understand about this hashing function better.

(Refer Slide Time: 16:25)



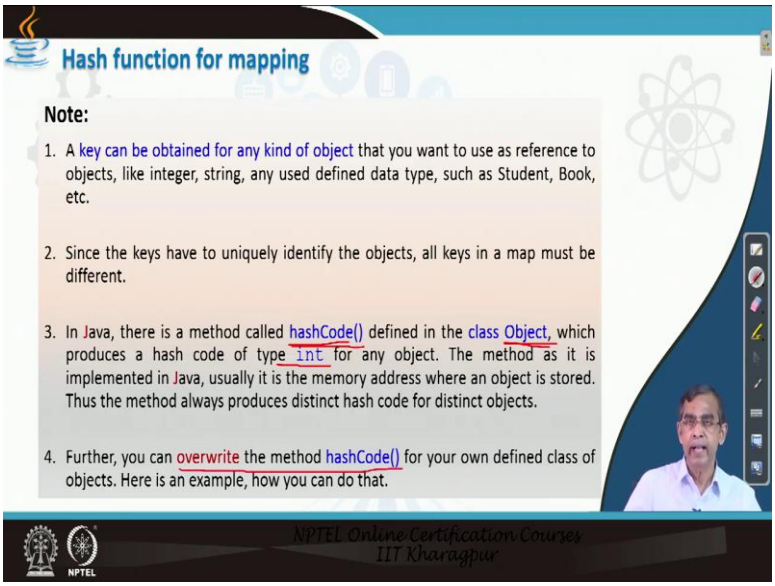

Now, I gave an example of a simple hashing function here is the hashing function let this hash function is defined like this so, you know modular remainder operation for any integer k assuming that k is an integer so, it keeps you, here is some example as we can see so, h 123 if we apply this model function it will be 5 so, hk is equals to 5 in this case, for another

example, h is 99 this gives you 4 and this one so, these are basically we can say address or memory where, the object whose key value k is stored. So, it is an object whose key value 123 it is another object whose key value is 99 it is another object is evaluates 2014.

Now, here judiciously you have to decide this key value so, that any collision if it is there it is avoidable or it can be avoided so, collision if it is there, it should not be (())17:26) it is not desirable for example, I can say suppose, h and 24 so, this is also equals to 5 that means, if there is another object whose key value is 24 they need basically gives the same mapping or memory location that means, the 2 objects whose key values 123 and 24 are located in the same memory location which is not possible, which is not desirable.

So, hashing mapping or hashing functions should be chosen in such a way that for every objects they are the memory location it will return is basically unique and that is the problem which basically needs to be handled very carefully so, this is also a very simple example but, these kind of examples may not work always so, you have to take care about this kind of problem and that is the only issue that we have to think about how this can be solved there.

(Refer Slide Time: 18:32)

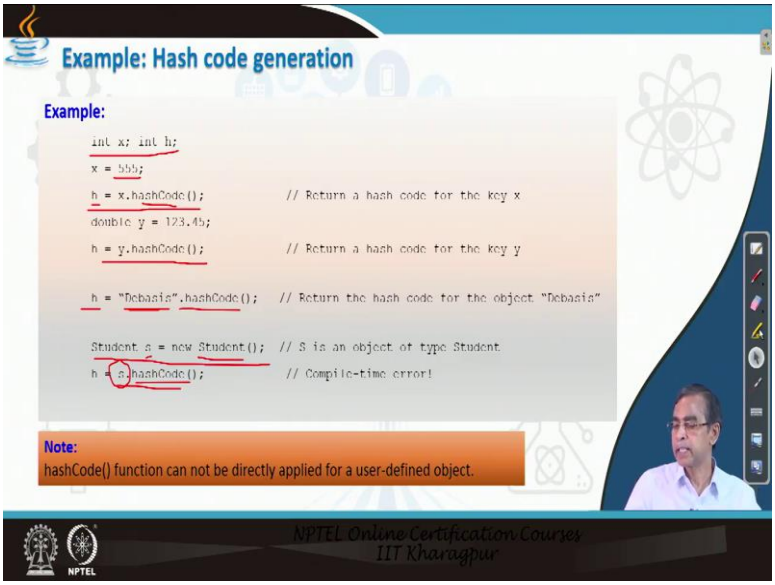



Now, in order to resolve these issues, there is a concept this concept is called hash code hash code is one method which is defined in the object class, the class object is a super class of all objects, which is defined in java dot lang package this means, that any objects that you define basically is a child of the parent class of objects so, it all objects that you define by any class declaration like say student, person, employee whatever it is, they are basically the child class of object so, hash code is defined in the object which is the default method and so, as any

object is a child class of object so, by virtue of inheritance this hash code method is also available to this one.

Now so, this hash code basically gives say an integer value maybe long value it basically essentially a memory address where a particular object is stored so, this hash code therefore is a very clever solution that is defined by java developer which basically if you give a key value for any object or set of key values to this hash function then, it will calculate a unique address for that object and however, you can also override the hash code method in your class declaration we will see exactly how such a overriding of method is possible or overriding the hash code that means, you can define your own hash code also and usually that is more desirable for the programming point of view.

(Refer Slide Time: 20:16)



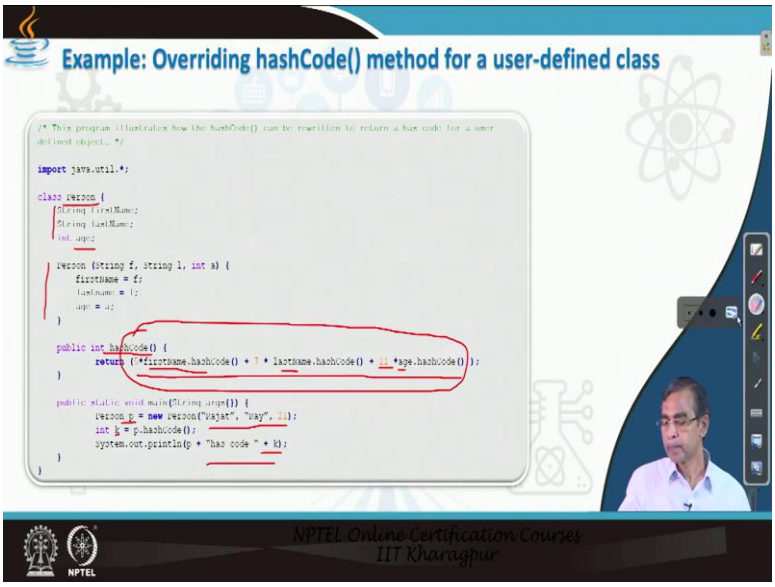

Now, let us consider an example for this so, here x and h are the two integer variables declared suppose x is 555 then, here h and then hash code if we call for this integer it basically return what is a hash value so, it basically it will return for some mechanism it is there, which mechanism? That is basically, it is internal to the java development we should not think about how they convert an integer value to hash value so, this hash value it will return you can write the programme and you can call this method for any variable or any objects and you can find it.

Now, here double y and if you call the hash code for this y, it will basically give a unique value corresponding to the hash code of y so, it is basically just a hx hy the hash mapping we can say now, for a given string also this is a hash code so, usually we use a password or login

id so, computer convert those password or login id into hash code and that hash code is basically a one way function given the hash code it is very difficult to come back to the original for example so, this is the string for which the hash code that h will return and from this given we will not be able to come to the what is our original input string for which the hash code is h.

Now, as another example, let us consider this so, h is an object of class student now, (())(21:55) hash code we can obtain in what is that hash code of this object h but, this is possible if we define our own hash code for this object otherwise called this code is basically sometimes give you a compile time error and it is avoidable because, it will try to give that default hash code which is basically compile time does not accept it so, here basically you need to define your own hash code for the newly defined class h for example here.

(Refer Slide Time: 22:27)




Now, let us see how such a hash code can be overridden and that is a programme you can consider this programme for an example, this programme is basically declare a class the name of the class is passion it has three fields, the one is first name last name and age and here is a constructor of this class to initialise objects and here basically we are overriding hash code how we are overriding hash code and this is the one method that we have decided, how we have decided it is up to the programmers ability.

So, here we decide that what is the hash code up string first name multiplied by 5 plus that hash code of string last name multiplied by 7 plus the hash code of integer h multiplied by 11

and then it will return a very complex on numbers which is basically considered the hash code of a particular object which basically we creates.

So, for example, here an object is created say p and having these are the different values for the objects and then it will return the hash code k for this object and you can print this as code if you print this hash code usually you can see it is in a hexadecimal string that it can return so, this way we can define the hash code of our own.

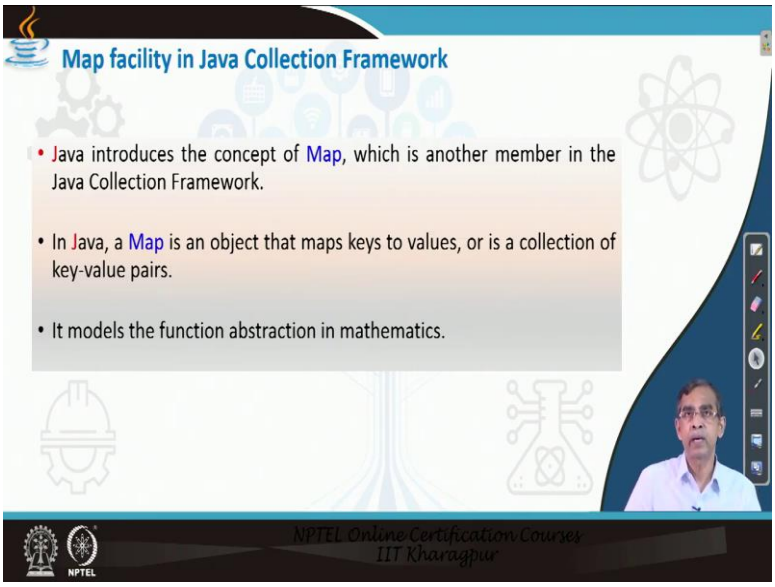
Now, here is basically is the intelligent aptitude that needs to be applied how we can define hash code for all objects, which should be unique for every objects that you have to deal with now so, there are certain theory behind it, those theories are you want to skip if you are interested, you can search many literature regarding has code generation you can find many ideas from there so, there is not a single idea rather many ideas to generate the hash code it is there.

(Refer Slide Time 24:30)


So, we have learned about the concept of hashing and hash code generation and how this hash code basically gives a unique representation for a given key value of an object where key value can be only a single attributes or it can be multiple attributes as we have planned in the last examples by our first name last name and age are considered as an attributes to generate the objects. So, whatever it is not an issue so, key value is not necessary that is a single value, it can be a composite value what I want to mention categorically.

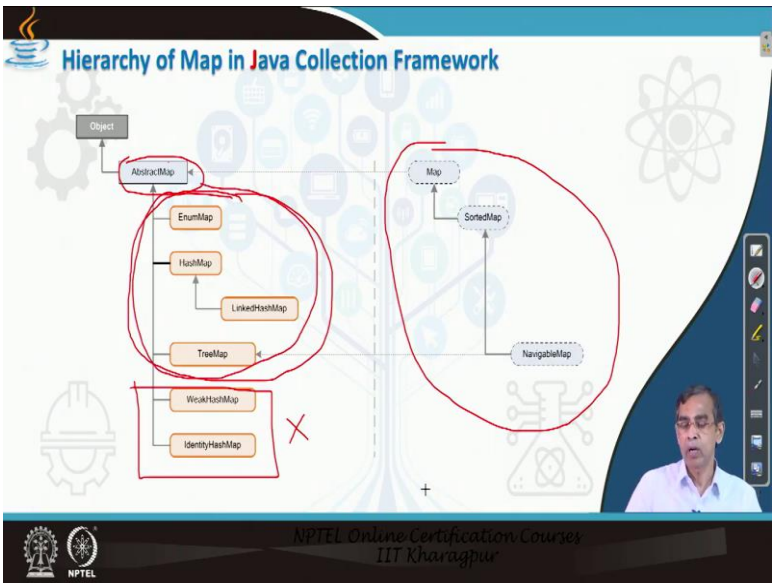
Now, let us come to the concept of so, these are the things that we have discussed about from the theoretical point of view in the conceptual level but, many things are made easy what the java developer has done for their programmers so, that we can have it so, that is the map framework for which we can have all those things automatically we do not have to bother about, how this hash code, how it can be generated, how it can be used, and whatever it is there now, let us see what is the map support which is there in the java framework

(Refer Slide Time: 25:28)




Now, java consists of two framework the collection framework and map framework, map framework is basically to deal with these tables are mapping that is there.

(Refer Slide Time: 25:41)



Now, the map framework if which is the class hierarchy that is there, it has this class hierarchy there are a number of interfaces defined in this framework, the interfaces are mentioned here, namely map, sorted map and the navigable map, those are the interfaces which basically have that design, any class you can design, if you want to implement a particular interface so, interface has the declaration of all methods or fields or any constant in them, there is no method defined there.

Now, so there are certain methods which basically you can use in your programme to create the object in order to manipulate your tables by virtue by means of your mapping mechanism the now important the classes which basically we consider are there now, there are two classes, these two classes are basically for the java developers use only so, we will not consider all those things in there.

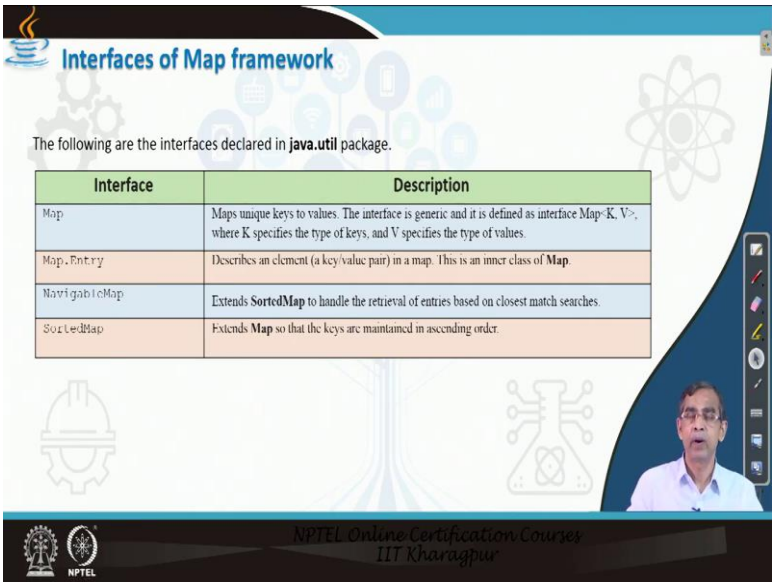
Now, so, there are four classes one is called inner map, hash map, linked hash map and Tmap this basically these classes allow you to create a collection this collection is basically table and automatically the indexing and everything will be taken care of by this now, using these classes you can create objects and for this object you can call many method how to add new object into this table, how to remove, how to access, how to search so, many operation those are usually require for the fastest access from the table.

Now, here again in Enum Map easy basically is extends the abstract map, abstract map is abstract class actually, that hash map is also extend abstract map linked hash map is basically extend the hash map and Tmap again extend the abstract map so, this means the again, you see abstract map basically implements map and then sorted map is basically extends the map, navigable map again extend the sort map so, all methods which are defined in map is basically a by means of inheritance also included there in addition to the methods in map they have their own method defined in sorted map and everything, all the methods which are defining these interface are implemented in all these classes.

So, these all these classes are implements all the methods which have did here so, in our subsequent few slides, we will try to understand about what are the interface competitions, what are the methods are there so, that we can use them in our programme regarding programming using map will discuss in the next video lectures today, we will just have the familiarity about the composition of this map interfaces and map classes it is too much also

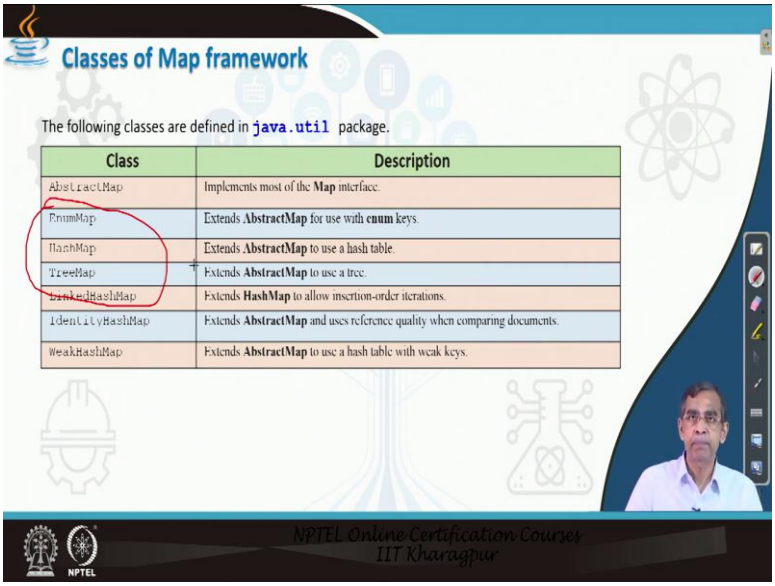
but, I will try to give a brief summary about all those things so, that we can have a quick look about this.

(Refer Slide Time: 28:30)

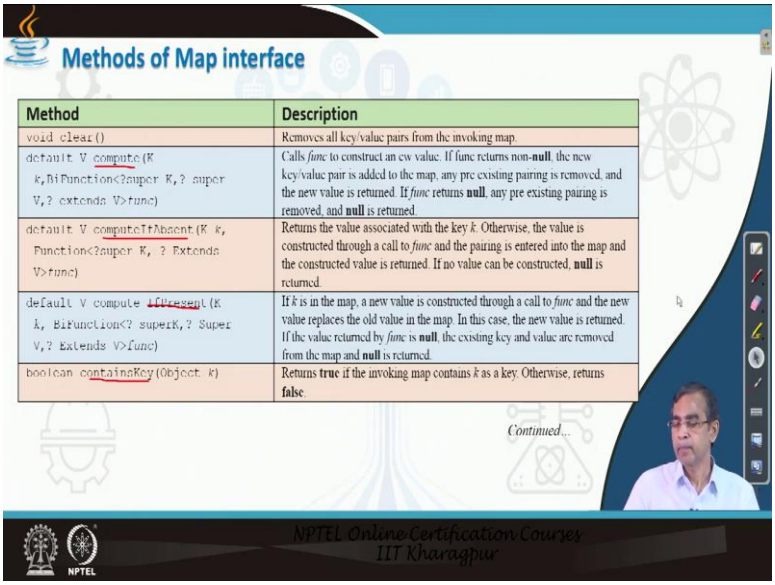

Now, let us have the map interfaces we have told you that, there are three map interfaces one is map another is sorted map and another is a navigable map in addition to this, there is also on another inner class is define it is called a map dot entry which is within the map interface it is there so, all these map interfaces basically is a generic or what is called a type, the type is of k and v, where k is all generic type and v is of an index so, k can be any type v can be any type and v can be any type of object k can be any type of object it is the concept is there, so, is the pair.

(Refer Slide Time: 29:14)




Now, there are so, these are the different interfaces that we have discussed in addition to the interfaces, there are many classes as you have already mentioned, these are the important classes which are basically used in programming namely Enum Map, the hash map, linked hash map and then Tmap.

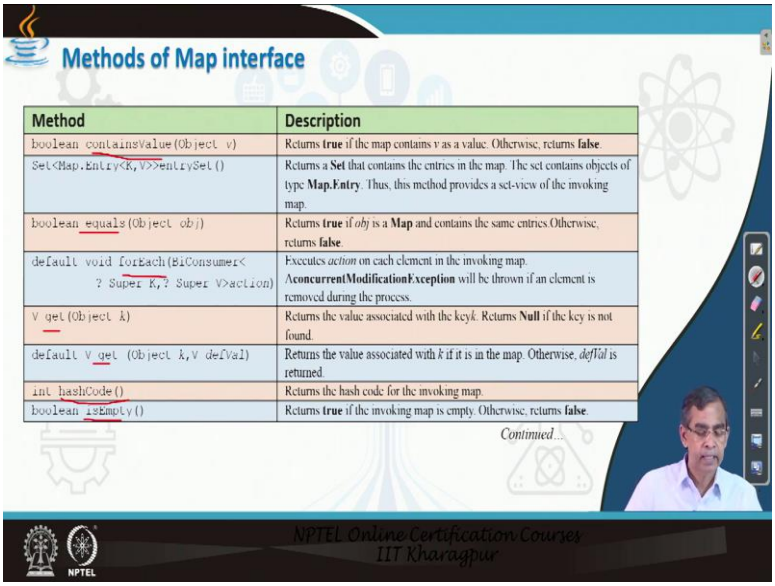
(Refer Slide Time: 29:38)




Now, let us have a quick tour to all interfaces that mean we will try to give an summary of what are the different methods are defined there and those methods are ultimately defined in the classes namely, Enum Map hash map, linked hash map, Tmap etc. So, we will just import familiarity will see, what are the different methods are the details discussion of the method, it is beyond the scope of this lecture and further detail discussion on you I should recommend you to check the oracle websites that is the document where the detailed documentation of each interfaces with examples are available.

So, there are many methods as we see here the compute method is basically how a key for a given key the hash value can be computed and is a compute if absent, there is a if present function and then contents key that mean whether, a key is available in the collection or not so, those are the methods are there.

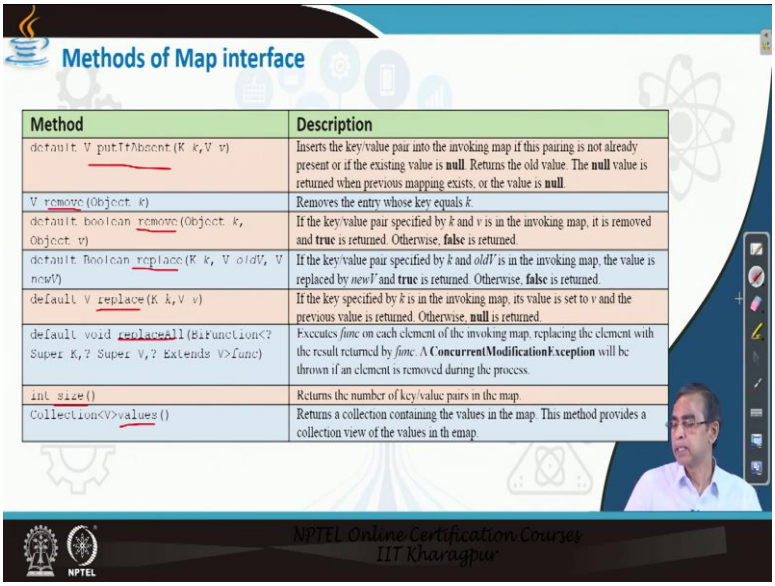
(Refer Slide Time: 30:41)




Now, there are a few more methods also it is defined there here, also I have included here contents value it is basically if a particular object is present in the table or not equals whether two objects are same or not for each basically for the traversal here get is a very important get means, we have to access an elements.

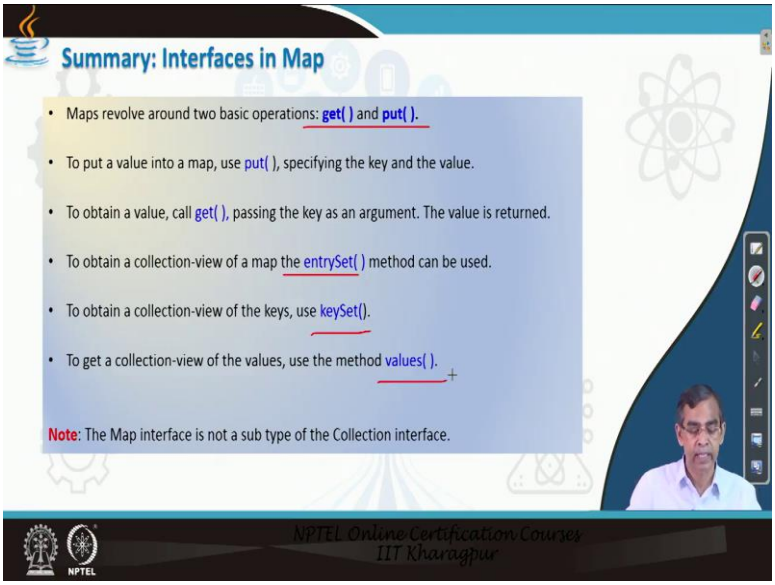
So, get method have the two different version that means, given a k whether, how we can get the object from the table and then, hash code is a default method it is basically there which is basically written the hash code and whether the collection is empty or not that needs to be checked. So, these are the different methods by which we can simply know about the collection, a collection is in this case is a table.

(Refer Slide Time: 31:29)



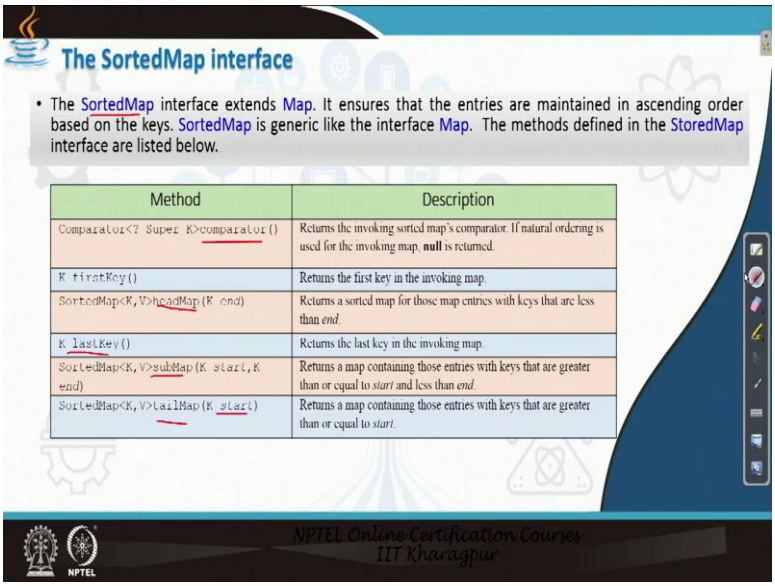

So, there are so, many methods which are defined there here, the put if absent the get method is basically to access it put is basically to store it so, there is a different method and then, put version is there and remove is basically to delete. So, there are many remove method is there replaces basically modification t hat means, if you want to modify certain key values or objects replace all if you want to modify all and then, size basically what is the size of that tables and then, value is basically all elements in the tables can be obtained using this method now, we will learn about the use of these methods while we discuss the programming.

(Refer Slide Time: 32:09)



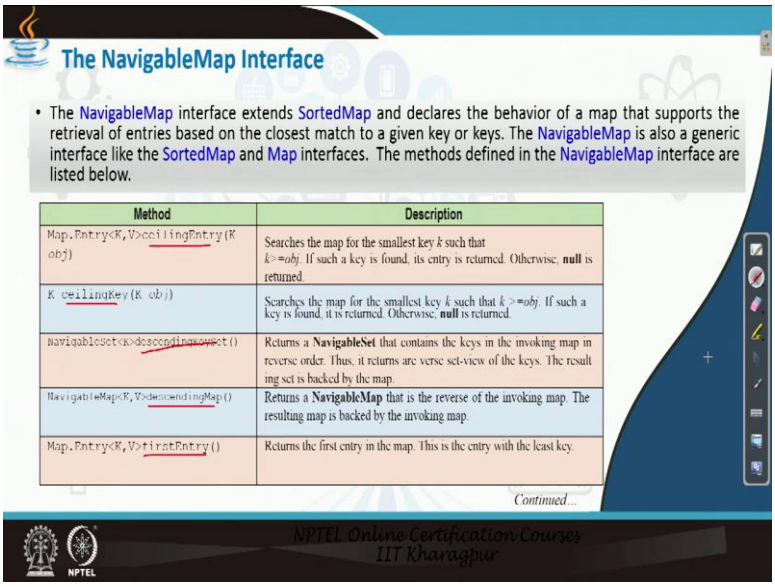
So, mainly so far the programme is concerned these are the methods we have to learn little bit sincerely so, that we can understand about it, when we will go for programming then, we will discuss about all these methods in details, the get and put method and then, entry set basically it will basically give you the set form of the collection or tables key set is basically set of all keys those are there in the table and values basically all elements those are there in the tables means, all records we can say so, these are the important methods basically we shall use in our programming or any programmer should use all those methods in their programming practice.

(Refer Slide Time: 32:49)



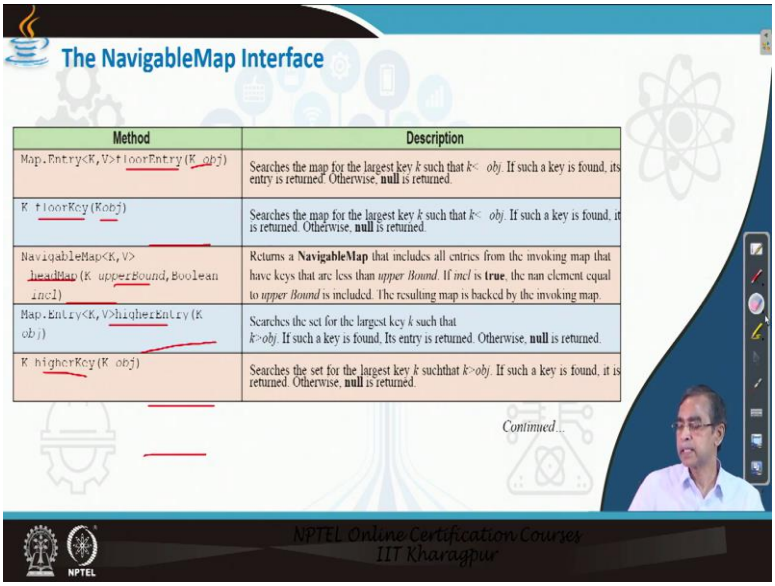

Now, so like this map interface, there is sorted map interface also similar to the interface there are many methods which are declared there for example, comparator, hate map this is the first elements the last element then, sub map subset of the table tail map the last few elements of the table given the starting values of key so many things are there and here sorted map is basically the difference between the map is that in case of map all the tables objects are not necessarily in sorted order but, in case of sorted map, it is basically sorted as for the key values of each objects so, this is the sorted map.

(Refer Slide Time: 33:32)


And navigable map is basically the idea about the closest key value if it is given there, if it does not match perfectly, then it will try to find the record or objects which are the closest match so, that is why the navigable map has been implemented there and it has many methods there is basically ceiling nt that means, it is some key values which is greater than some key value which is basically a ceiling key, ceiling entry and descending key set that mean it will give the set of values in descending order descending map it is all objects in descending order fast entry which is a fast object it is there.

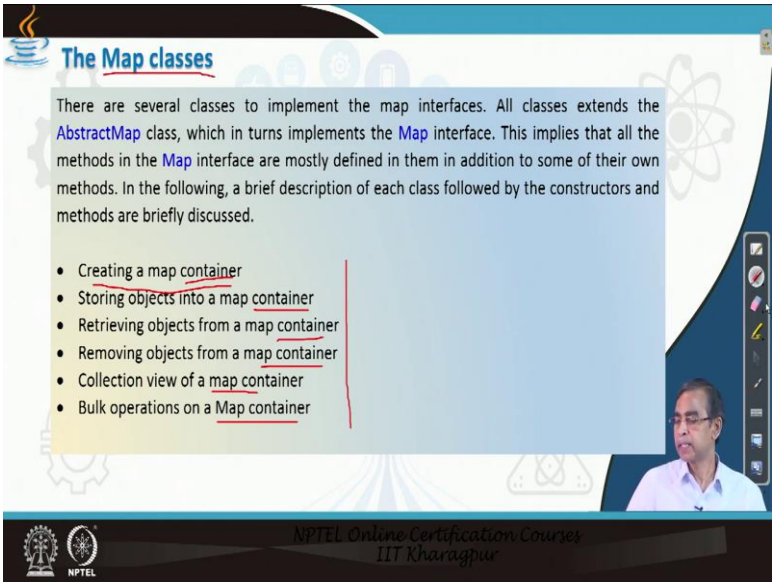
(Refer Slide Time: 34:13)




There are a few more methods also it is there like that floor entry as it is opposite to the ceiling entry below to a certain values the key and then floor key is all the objects all the key values basically certain value layer, which is basically the object that is there means, below objects or is basically a lower bound upper bound like so, there is also hate map giving the upper bound and then Boolean that means, whether will include all or not then higher entry higher key.

So, those are the methods basically we will discuss whenever we discuss programming in details, illustrating how these methods work, and then you will be able to understand but, this is for the summary only these are the methods have available for you so, that you can use it.

(Refer Slide Time: 35:03)

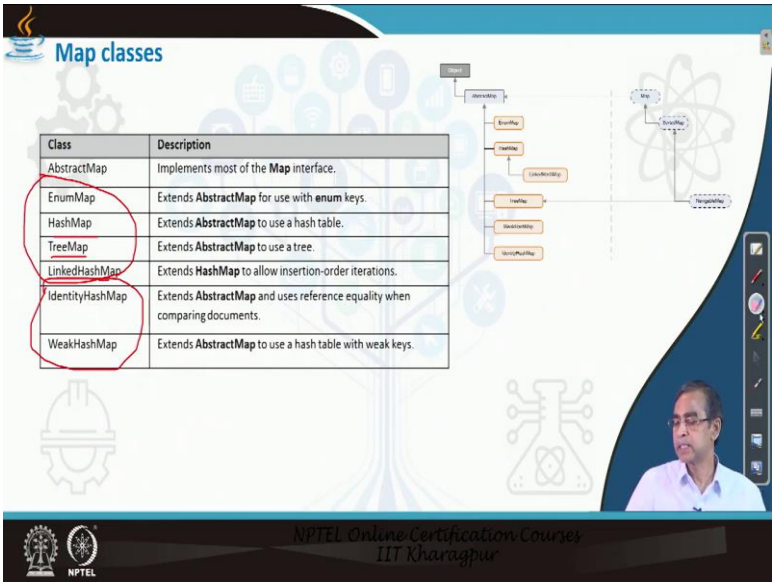



And there is so, in these are the different methods and likewise the map classes are there, the map classes are ultimately that you can create an object that is a collection object in your programme and then, it basically you can manipulate this programme and again I repeat a collection according to this map is just a table a collection according to the java collection framework that we have discussed about vector is a collection, array is a collection, array list is a collection, T is a collection, linked list is a collection those are different collection that we have discussed but, here all these map basically gives a new form of collection, this is called a tables actually and the table collection is characterised two sets on the other hand, all these collection are characters only one type generic type so, there is that difference.

And now, so far this map classes is concerned so, how a map classes it is also called container, how a collection or it is call rather we can say the container, a container can be created, how container can be sorted, how some objects in a container can be retrieved, how some objects in a container can be removed, how some container can be viewed and how the bulk operation copy from a set to another set all these things can be done or a table can be copied from another table or two table can be marched and so many things are there.

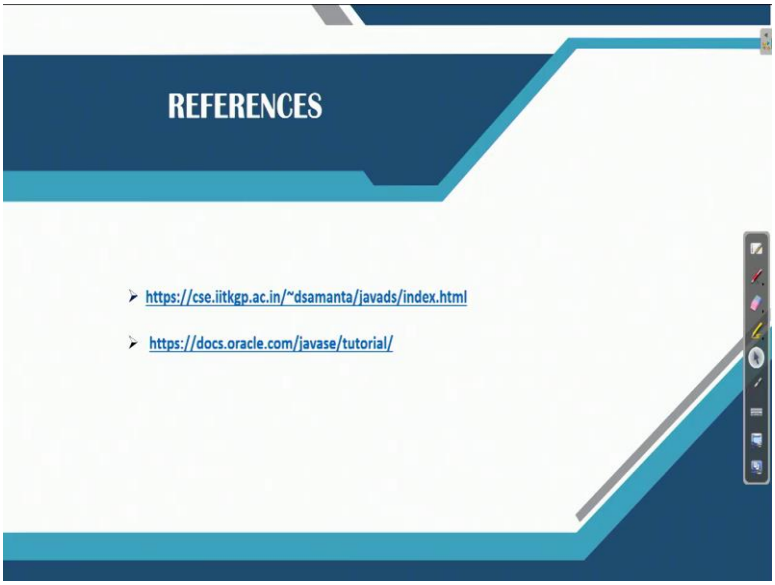
So, these are the basically activities or operations that we will be able to achieve using the different classes which is defined in the map framework and these are called a map classes and so far the different map classes are concerned I have already mentioned that four map classes are there.

(Refer Slide Time: 36:46)




So, those map classes are basically in Enum Map, the hash map, Tmap, linked hash map these two maps are not used for the programming purpose so they are not included in our discussion so, this is basically overall the map framework. Now, there are different methods truly the methods those are already learned about interfaces they have been defined all these interfaces so, there is no extra method in each classes actually to be studied further so, however only thing that needs to be understood about that, how the different collections or container objects can be created so, constructor needs to be there.

(Refer Slide Time: 37:25)



So, in our few, in our next few slides, we will try to understand about how the different constructors can be created that will be discussed in our next video lectures, where we will discuss about the different classes and the different constructors, which are basically used to create the objects and using those objects we can manipulate different tables. Thank you.