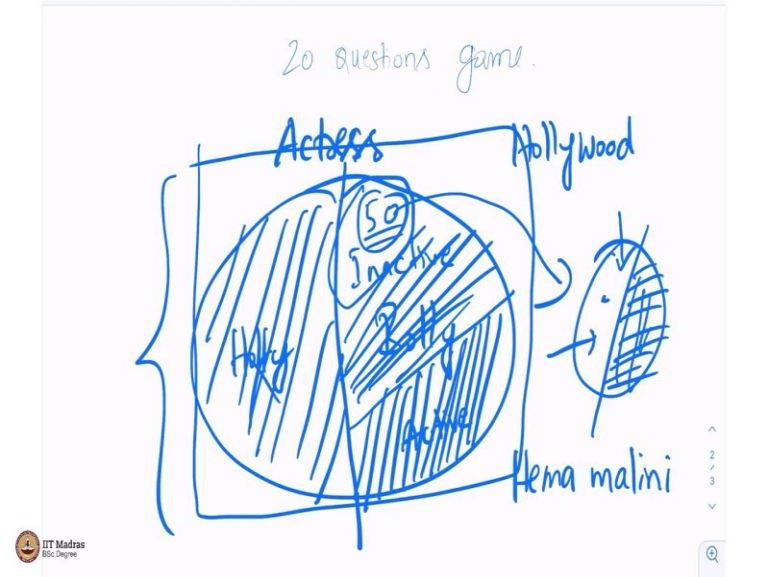

# IIT Madras

## ONLINE DEGREE

**Programming in Python**
**Professor Sudarshan Iyengar**
**Department of Computer Science & Engineering,**
**Indian Institute of Technology Ropar**
**Mr. Omkar Joshi**
**Course instructor**
**Indian Institute of Technology Madras**
**Online Degree Programme**
**Introduction to Binary Search**

(Refer Slide Time: 00:16)




So, we all have played what is called the 20 questions game. If you did not know, let me illustrate that with an example. It is called the 20 questions game. The point is, you will have something in your mind and the opponent should guess what you have in your mind. Now, I will say, I have an actress in my mind. And you must guess who is this actress.

I have an actress in my mind. Who is this actress? Your first question would be, there are so many possible actresses in the world. Who am I talking about? You should guess that person which I have kept in my mind. First you will ask, is the actress from Hollywood or not? And I will say, no, she is not from Hollywood. Hollywood, no. Assume that the actress either is in Hollywood or Bollywood then I am reducing the sample space into half because this side was Holly, this side was Bollywood and you strike this off. Half the possibilities are gone.

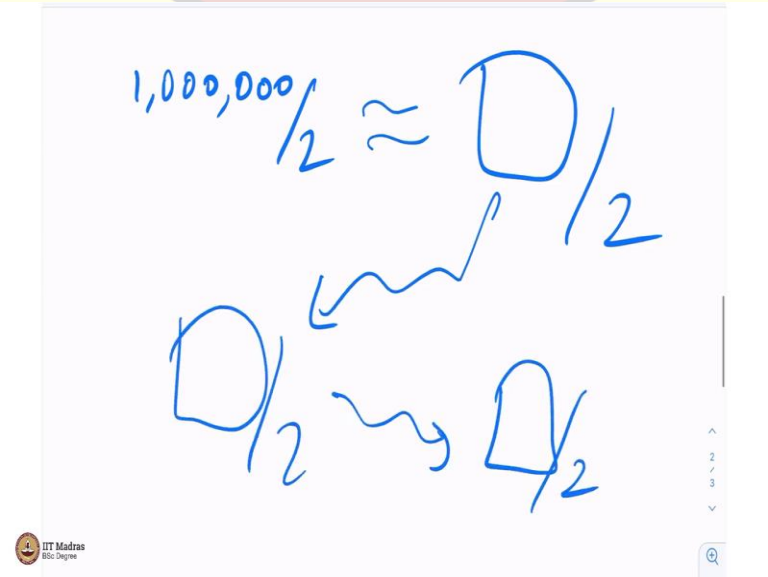
And next in Bollywood, you ask this question. Is the actress still acting? And I will say no. And then that is roughly a lot, active and inactive. And active is again gone. The actress is not really acting today. And then you will ask this question. Is the actress about 50 years of age or less than 50 years of age? I will say she is above 50. Again, this much is gone.

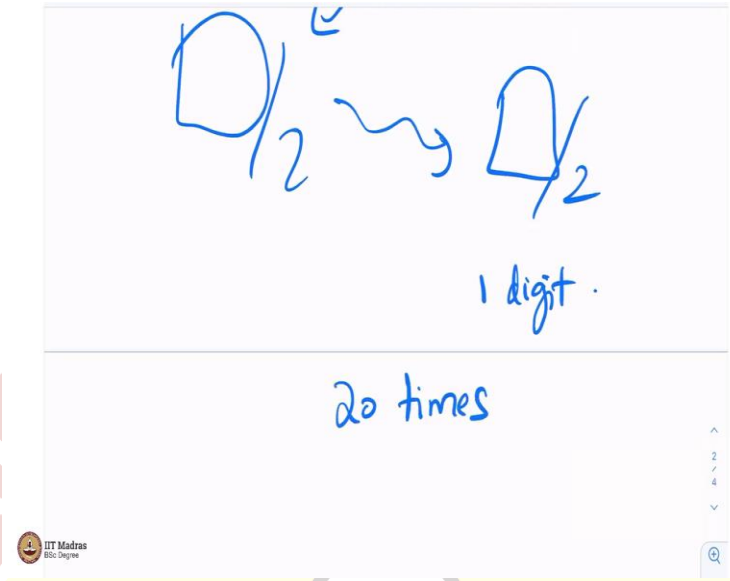
You see the sample space is reducing like anything. In a few more questions, you will just narrow down the actress name very quickly. Your next question possibly would be, is the actress into politics? And I say yes, she is into politics. And then again, this little thing sample space, let me just write this down, very small sample space, in politics, not in politics.

Not in politics gets wiped out. And then in politics, an actress from Bollywood who is not active anymore in movies is into the politics. How many such people are there above 50 years of age, very, very few. And you will start guessing is it Jaya Bhaduri, is it Hema Malini. The moment you say Hema Malini I will say bingo, you got it right.

So, out of roughly a million possible actresses, you narrowed it down to Hema Malini by cutting down the sample space by half every time you were asking this question. Roughly half, it may not really be half. But the point is you sort of squeeze the sample space, this space, all possibilities is called the sample space. You try cutting it short and try making it very, very small.

(Refer Slide Time: 03:58)

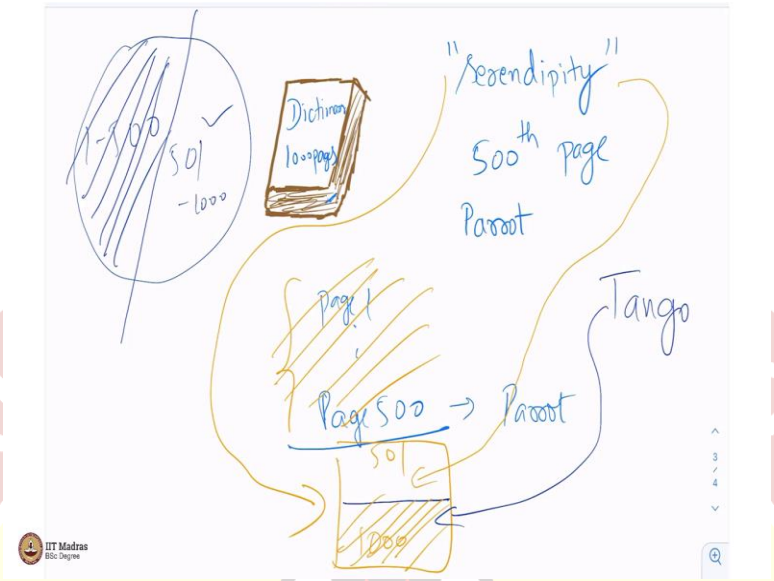


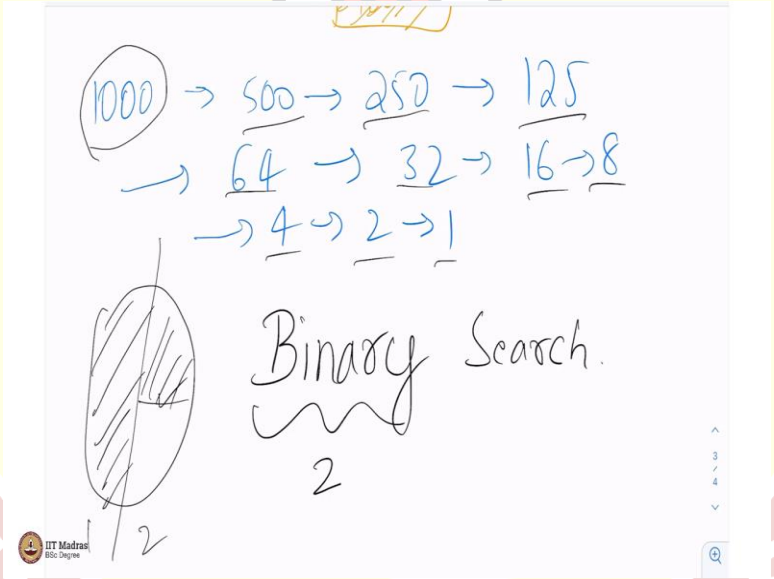


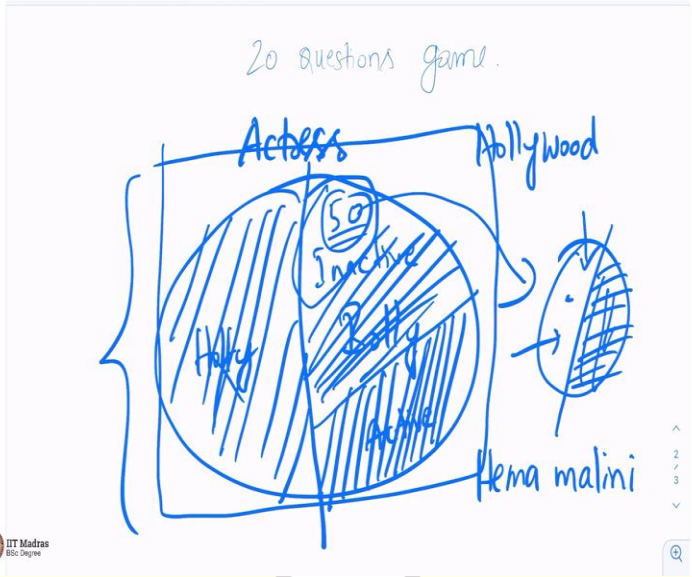

Let me take another example. You take a very, very, very big number. How big, let us say a 1 million, divide it by 2, whatever you get, divide that by 2, whatever you get, divide that by 2, whatever you get, divide that by 2, how many times should you do this for that to become a single digit, one digit number? Simply 20 times, try doing that.

Take 1 million divide it by 2, divide that by 2, that by 2, that by 2, and so on, in less than 20 times, you will see that you have in fact exhausted the number and made it a single digit number. Now, how is this happening? So, let us look at a third example now, 20 times will make this number go very small. Although the number of actresses too many numbers, you can still narrow it down by halving. This is called, what is that, in English, it is called halving, half.

(Refer Slide Time: 05:17)






So, let me take yet another example. This example is going to be a very obvious one which is, let us say you have a book with you. Let me use my writing skills to, drawing skills which is actually very bad, to write a book here and then, let us say, here is a book. And what I will do is, this is a dictionary. This is a dictionary with around 1,000 pages. And you are after searching for the word, some words that you will not know.

Maybe you wanted to check the meaning of the word, what is a complicated word which we may not know, let us say serendipity. You want to search this. What you do is the, dictionary has 1000 pages, you open some random page. Let us say you opened the 500th page. And this 500th page, first letter is P, the first word is P there, here, half of it, which means the first 500 from page 1 to page 500. 500th page had parrot.

So, your serendipity word should be in the second half of the book between 501 to 1000th page. It cannot be here. It must be here only. So, what you do is you discard. You do not worry about the first 500 pages. So, even in a dictionary the sample space here meaning all possible places where you may get the word serendipity is 1000 pages, but then you cut it by half. What you say is you cut it by half.

First 500, 1 to 500, you say the word is not here. The word is in 501th page to the 1,000th page search here. And then again, you will open a page and you will reach t for, there is a word from t it is called let us say Tango, the word Tango. And you know very well that Tango is let us say in the 750th page here. Your word serendipity is definitely then going to be between 501 to 750.

This can also be discarded, because 750th page became Tango. Every time you are cutting it by half, you go ahead like this. And exactly the way you, when you cut the number by half, it reduces drastically. A dictionary with 1000 pages it is getting half and half and half very soon you should get the word Tango. 1000 pages become roughly 500 pages, roughly becomes 250, roughly becomes 125 and then 64 roughly, of course, then 32, 16, 8, 4, 2, 1.

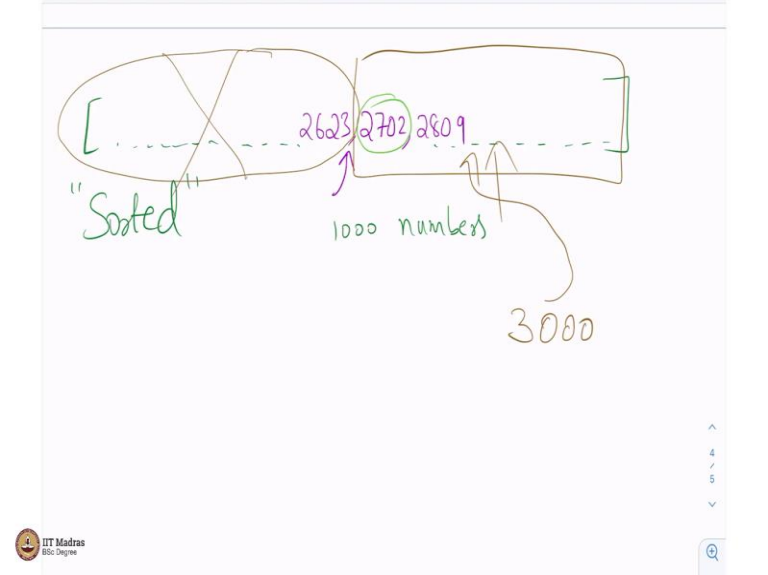
Instead of checking the entire dictionary for the word serendipity, you check the middle one and then see what the world is based on that you discard half the dictionary and so on. This is a popular method called the binary search. Binary means two. Binary search which means you check, you divide the entire space where you are searching into two parts and you discard one half of it. It is always about two parts.

And then the part where you suspect the search element is, again you cut that by half and then you say it is probably here, so on and so forth. And as you know, very quickly, instead of searching for all 1000 pages, you cut it short 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, in 10 steps in a 1000 page dictionary, you get to find the word that you are the lookout for. This is popularly called the binary search.

I am sure most of you already know of it. If you do not know of it, you have used it in many works of your life. We use this almost always in our lives. You have used a dictionary, you use this method. Anything that is arranged in alphabetical order there is to be what is called a telephone directories initially now nobody uses it that would again be a very big book of all possible names in the city. And you can find out a person's name in that by going alphabetically and what we do there is binary search.

In fact, the whole of Sherlock Holmes stories if you observe, the idea there, in any detective stories for that matter, is to narrow down the possibilities and converge to very, very few possibilities, just the way I did for what was that the Hema Malini actress that I had in my mind, by asking the right questions, you can easily narrow it down to the right answer. So, detectives use this, dictionary searching uses this, the 20 question games uses this and then a simple dictionary, telephone directory or even having a number you use this method.

(Refer Slide Time: 11:51)




And finally, we come to this question of there is a list, a big list with around 1,000 numbers and then there is a big list and it is sorted. Please note it is sorted. Do not forget this. It is sorted. And then my question is, in the middle, let us say in the, roughly in the 500th element, if you see, elements are 2623 and 2702 and 2809, you see now this number and you want to search for whether this contains the number 3,000, this list, you want to check if this list contains the number 3,000. How will you find that out?

What you will do is come to the middle which is 2702 to the left of it is a number less than that because it is sorted, to the right of it is a number bigger than that because it is sorted. And you know very well 3,000 if it exists, it should be in this part of the array and not in this part of the array. So, what you do is you discard this part.

You do not check that. You only check if 3,000 is there or not in this part of the array and this continues. You take in, go to the middle of this and check what is the number. The number is greater than 3,000, then it is in this part 3,000, if it exists, otherwise it is in this part. If you are finding it confusing what I will say, I am simply saying exactly what we have discussed here, same thing in terms of searching in a list.

You search in a list using the binary search technique. You, half the, you create half of the list and such there and then you go, keep on creating half and half and half of it until you see whether the number is there or not. So, let me try writing a program for us and see how we can

code to find an element in the list and how is this a very efficient method. I leave and explain what one means by efficient a method. Let us get into the programming details of this algorithm now.