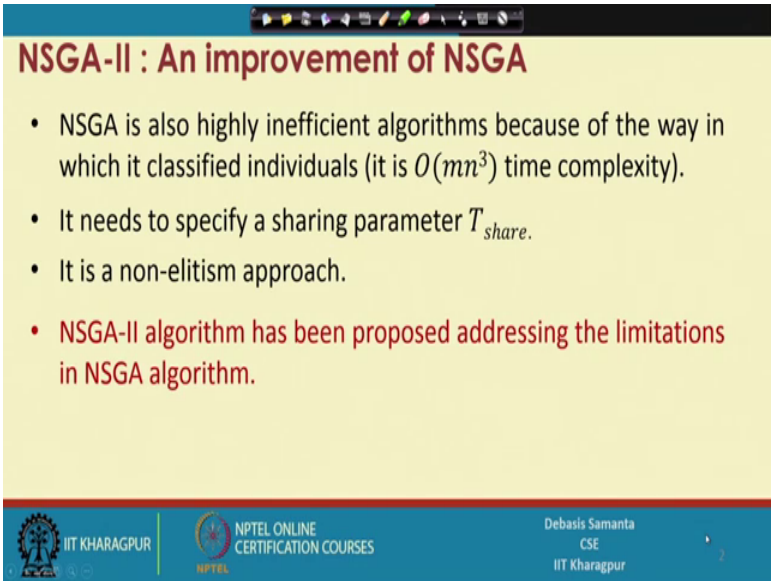
Introduction to Soft Computing
Prof. Debasis Samanta
Department of Computer Science & Engineering
Indian Institute of Technology, Kharagpur

Lecture - 32
Pareto-based approach to solve MOOPs (Contd.)

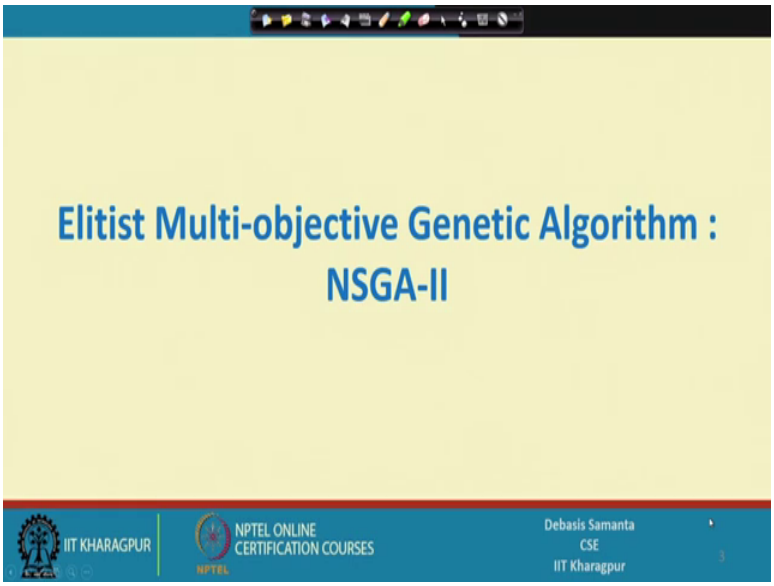
In the last lecture, we have learnt about NSGA non-dominated sorting genetic algorithm proposed by N Srinivas and K Deb. And we have discussed some limitations there addressing all those limitations the same authors with some other more researchers.

(Refer Slide Time: 00:38)



They have proposed the another algorithm, the next version, the improved version of NSGA, it is called the NSGA 2 algorithm.

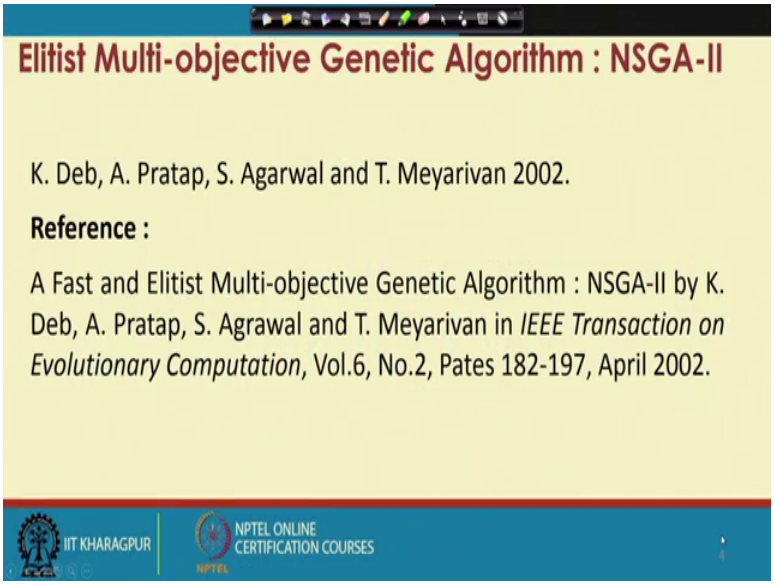
(Refer Slide Time: 00:48)



So, this algorithm is called elitist. Elitist multi objective genetic algorithm and abbreviated as NSGA 2, and the basic concept in the NSGA or rationale in the NSGA 2 is basically so that it can be compositionally much efficiencies, and then it gives better result compare with NSGA.

So, this is the idea about the here in NSGA 2 approach.

(Refer Slide Time: 01:19)



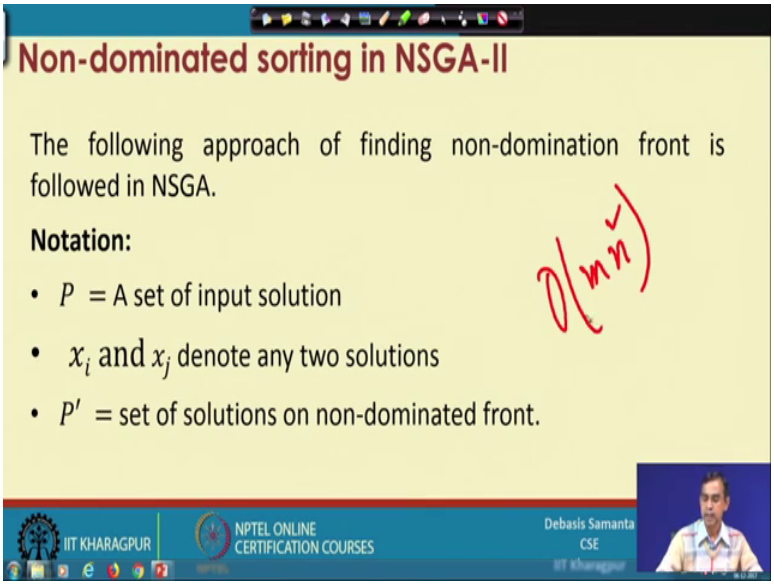
And here are the contributor who developed this algorithm, the contributors are K Deb, A Pratap, S Agarwal and Meyarivan. This algorithm NSGA was first time proposed

introduced in 1994, and around 8 years later this algorithm NSGA 2 has been proposed by the same group of researchers in the same lab. So, these lab is basically the lab they are in IIT Kanpur ok, and this lab is famous as NSGA lab.

Now, the work NSGA 2, first time in his published in IEEE transactions on evolutionary computation. And the title of the paper was a fast and elitist tist multi objective genetic algorithm NSGA 2. So, basic idea they claim that it is fast, and elitist. Now we will see exactly how the time complexity have been owned here, and then how the concept of elitists has been enjoyed here or is applied here rather, and the algorithm it is there as the algorithm is bit complex and lot of steps are there. So, we will discuss it in 2, I mean sessions, in the first time we will discussed the basic concept or overview of the things. And it considers some unique approaches there it is called the crowding sort. So, for the selection is concerned we will discuss in the next slides.

Now, let us start about the concept it is there in this algorithm.

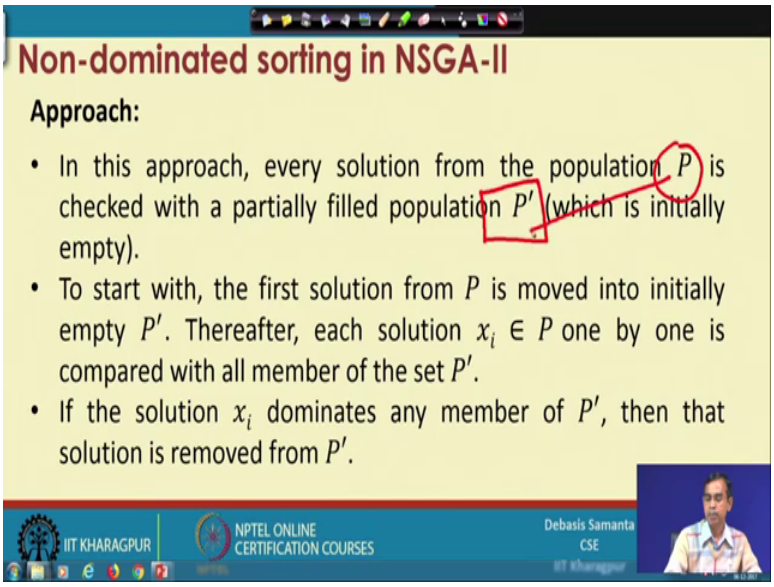
(Refer Slide Time: 02:54)



Now in order to understand we can recall that the NSGA basically consider the that a front. They basically calculate all non-domination front it is called, and then once the front is calculated, then they assign fitness values to all the solution in a particular front this basically the dummy fitness values, and then sharing the fitness value by means of needs count.

However, the procedure that it follows it totally different. This is because finding the non-domination front itself is a computation very much x expensive, we have discussed about that the complexity of finding the non-domination front is in the order of it is order of m N square, where m is the number of objective and N is the population size.

(Refer Slide Time: 03:40)



And we will discuss we will see that how the, it also considered the finding non-dimension front, but in a different way, we will discuss about this procedure.

Let us first consider the different notation that it follows to discuss about it. So, like the NSGA procedure, we consider the P is basically input solution set; that means, set of all solution belongs to a current population. And we denote x I and xa are any 2 solutions in the set of solutions, and here P dash is the set of solutions on non-dominated front. So, P dash be there any solution it is there.

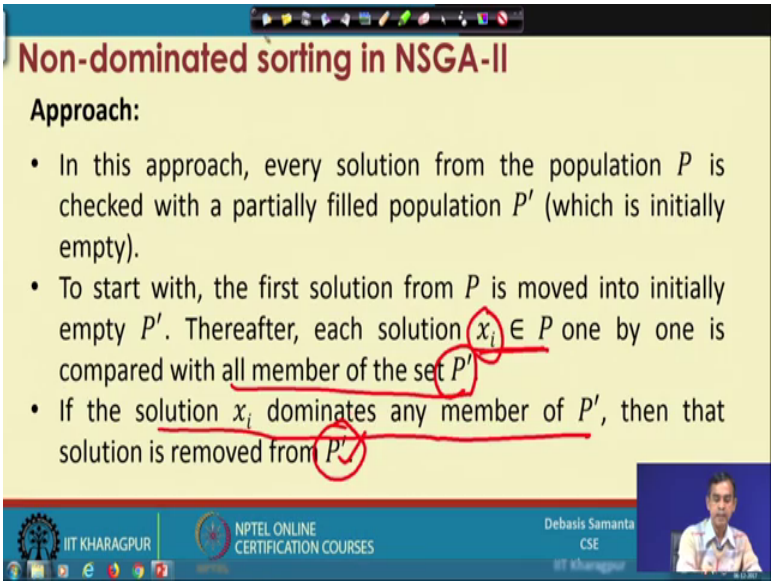
Now so, it is the, these are the terminal that will follow, and then the procedure the approaches is like this. So, here every solution from the current population said that is P, it checked with a partially filled population P dash which is initially empty. So, the idea is that given a set of solution P, we have to first find a non-dominated front P dash.

So, it is basically idea so, how the from the set of solution P we can find P dash? So, basic idea is that initially P dash is empty, and we can randomly choose ok; so, basically initially P dash is empty, and then we have to fill the P dash with some solutions they are

basically the solutions are in the domination front. So, P dash basically say concerned contains all the solutions belong to the domination front at any instant.

Now, to start with this; that means, filling the P dash from the P what basically idea is that we have to randomly choose on solution x i which is in the P one by one is basically; that means, all the solutions should be checked one by one which is there in the P.

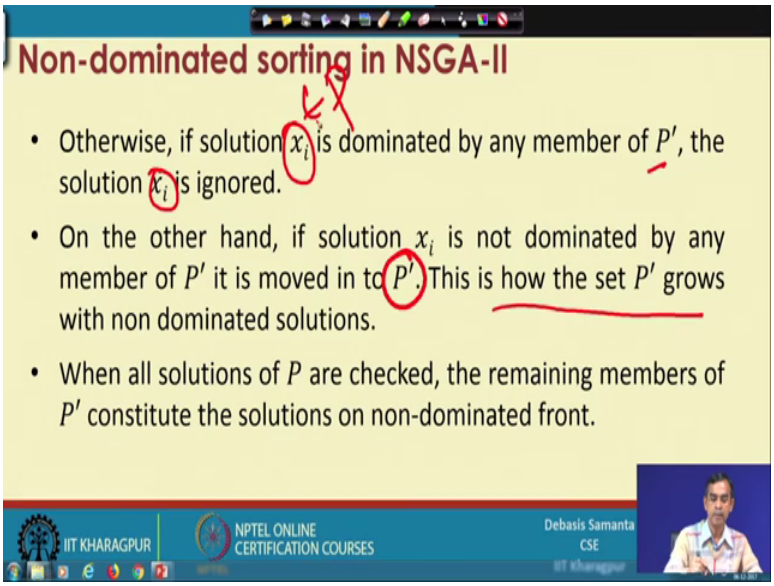
(Refer Slide Time: 05:42)



Now here so, so, compared with then this x i is compared with all members of the set P dash. So, initially the P dash is empty; this means that, x I should be there in the P dash initially. Or in other words, suppose P dash is not empty, then it basically check this on. If x is dominates, any member of P dash then that solution is removed from the P dash. The because the P dash should contents all solution which is not dominated by any other solution. That is why P dash should be a non-domination front.

So, it is removed there and the same procedure is repeated for all other solutions one by one, and then P dash will be gradually filled, and then P dash contents all the solution in the non-domination front.

(Refer Slide Time: 06:40)



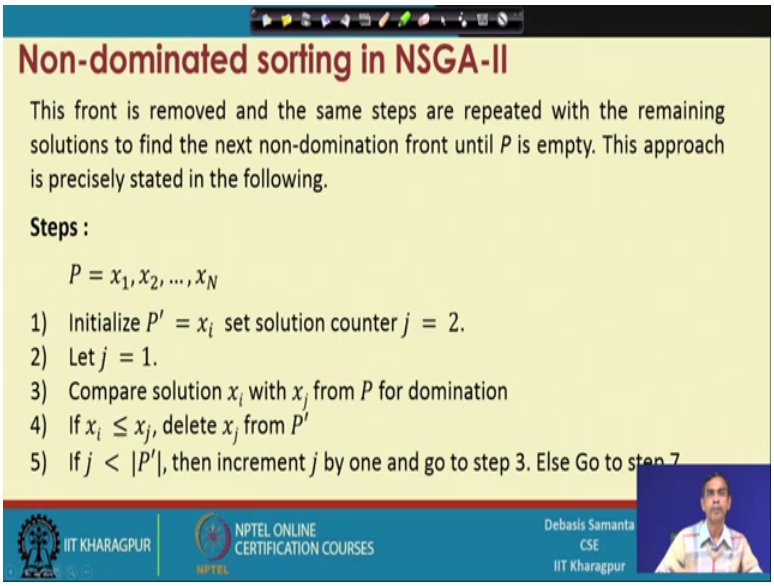
So, this is the idea about this one. So, here again repeated is if solution x i is dominated by any member of P dash, the solution x i is ignored. So, the solution needs to be need not be consider in the P dash.

On the other hand, if solution x i is not dominated by any member of P dash it is included into the P dash, because x i is therefore, the solution to be in the non-dominated form. So, this way the P dash grows when we checked all the solution x I belongs to P the current solution one by one. And when all the solutions of P are checked, then the remaining members of P dash constitute the solution on non-dominated front.

Now, you can understand this basically the technique in the last slide and these slides includes and this basically the clever idea about the previous approach. So, this solution is better this solution is basically find the non-domination front P dash.

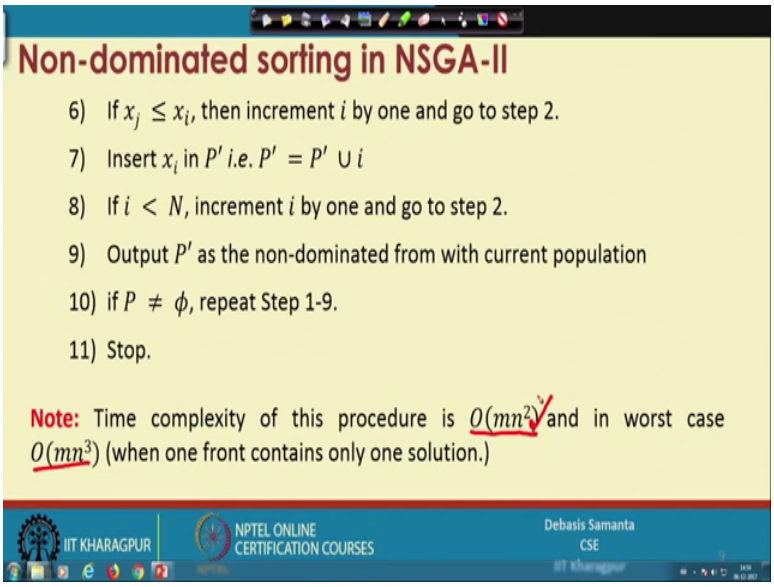
Once the P dash one, first front is obtain you will remove all the solution from the P dash, and repeat the same procedure finding the next front and so on. So, this method is therefore, different than the non-dominated sorting procedure in NSGA. And it is more faster than the procedure that is there in NSGA. So, this way it is a first method of finding non-dominated sorting front.

(Refer Slide Time: 08:07)



Now so, the idea it is like this so, so, it is basically it will these are the steps that we can follow, and more precise steps that is there in the calculation of non-dominated front there.

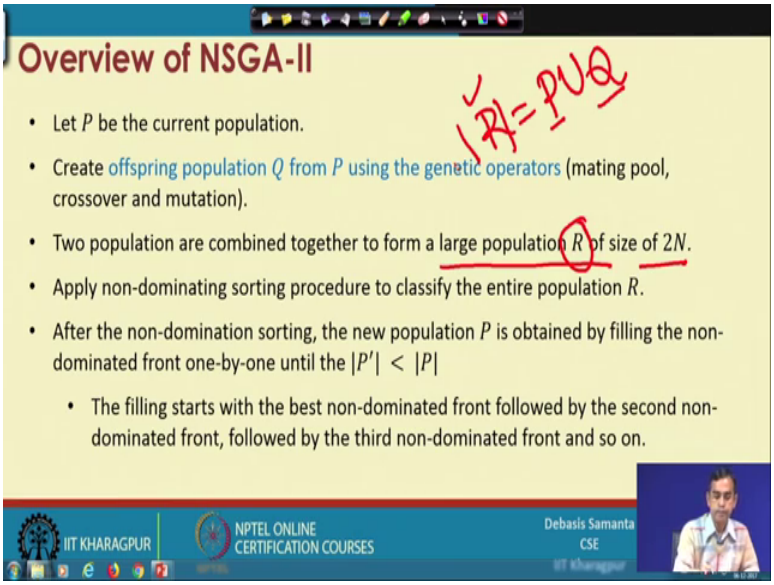
(Refer Slide Time: 08:22)



And we can observe that a detail calculation can let you that the complexity of this procedure is order of mn square, and in worst case it is order of mn cube, although the same complexity here, but still it gives the better what is called the time than the

previous one. So, so this way this is the improvement over the nsgo so far, the non-dominated sorting front calculation is conserved.

(Refer Slide Time: 08:58)



Now we will discuss about the basic approaches of NSGA 2. First the thing is that how to find a non-dominated sorting front. Now we have learned about that complexity of the 2 algorithms are same. So, anyway both NSGA and NSGA 2 needs to calculate all the front it is there. So, whether the new method that we have discussed just now can be followed or the method that is there in NSGA also can be followed in order to calculate the non-domination front anyway. So, non-domination front calculation is there in both the algorithm NSGA and NSGA 2, either say non-domination sorting procedure, or the revised procedure that is there in NSGA 2 algorithm.

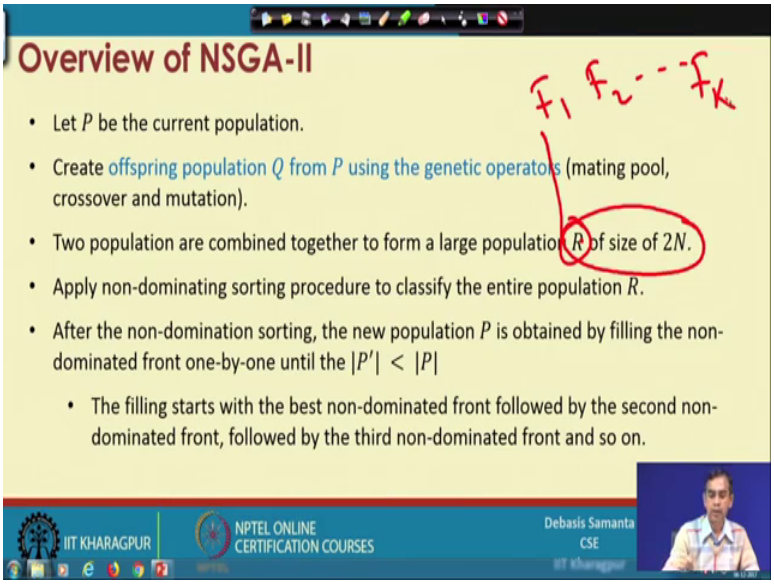
Now, let us see what are the basic step that it is followed and NSGA 2 algorithm, after the calculation of non-dominated sorting front is known to us. So, here is the procedure that I can give in a stepwise manner. So, let P be the current population; that means, the initial so, the current solution or the current generation. Then here the idea is that it basically from the current population P, it generates offspring population Q, right.

And the so, this is basically a reproduction; that means, considering the P, and whatever the method that is known to us in the simple genetic algorithm; that means, a how to create a mating pool the crossover and mutation, if we follow, then from the set P we can deny it another set Q. Basically Q is the next generation like.

So, basically from the current generation P, we shall be able to generate the new solution sets Q based on the reproduction techniques the usual. Now once the P and Q are known what the technique in NSGA 2 is that we combine all the solutions both from P and Q together a solution sets. That mean a P is the initial solution and then Q is the next solutions, then the combining the 2 solution gives a solution size or a population size let it be R is the size of 2N. So, R is basically all the solution those are there P and then Q. So, as P size of N Q is size N. So, the size of R is equals to N.

So, basically, we will developed a large population set of size 2N.

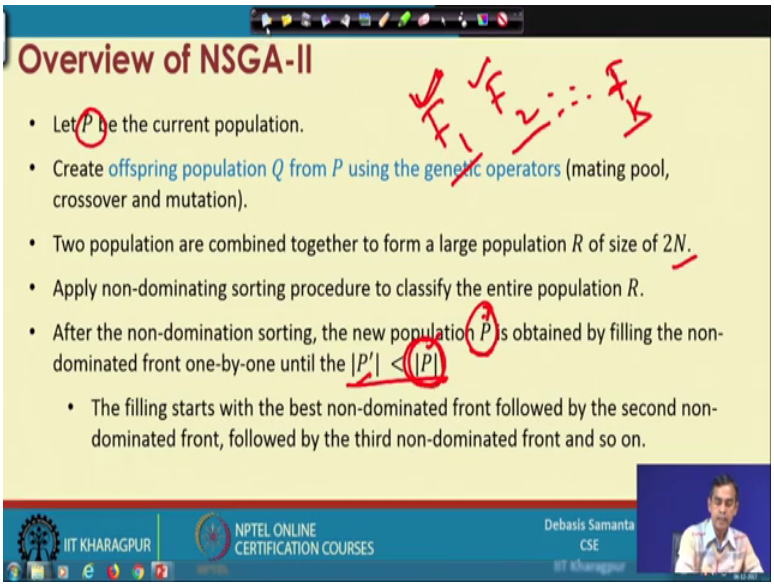
(Refer Slide Time: 11:47)



And then the on this solution set R we have to apply the non-dominated sorting procedure that we have discussed so that all the fronts can be calculated so, from R we will calculate all fronts. So, F 1 F 2 dot dot dot dot F k number of front say. So, k number of fronts that can be calculated using the non-dominated sorting procedure like.

Now so, after the non-domination sorting is done, the new population P is obtained now. So, this is the current population from this current population the new population P will be obtained by filling the non-dominated front one by one, until the P dash is less than P. So, here basically the idea is that so, suppose F 1 F 2, then dot dot Fk are the different front.

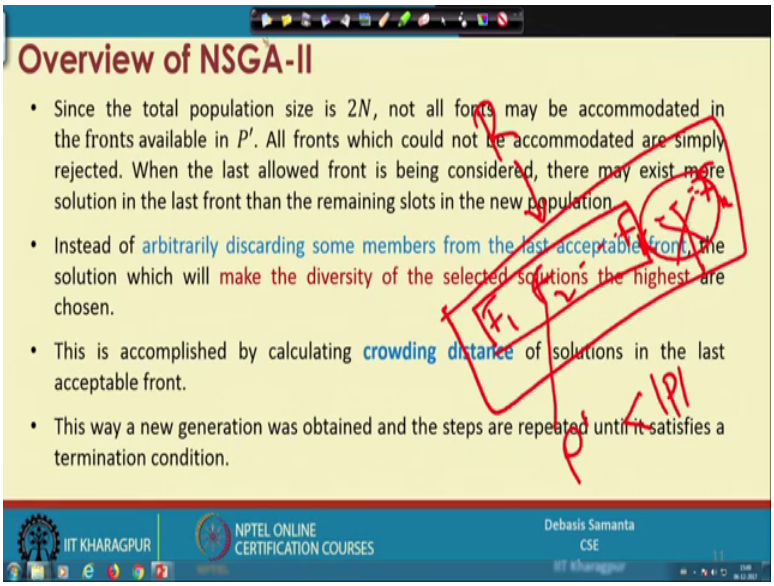
(Refer Slide Time: 12:40)



So, one p so, from the P dash so, from the P the next P dash has to be obtained ok. So, it is basically we are we are in the process of getting, the next population set P from the current population P. So, basically, we will select F 1, if we see that size of the next population is less than the population size that is basically N. So, if we select F 1 F 2 and we will go on selecting until this condition is satisfied. And you can say the last front is the one important front, basically, this front if we include into the current population, the lump the limit that is the size of the population will exceeds.

So, that last front is the one important front that needs to be taken into care and the special treatment to be applied. Now we will see exactly what is the special treatment that can be applied to the last front.

(Refer Slide Time: 13:55)



And here is basically the concept of concept of some selection. And that selection is a vital there in NSGA 2 and is basically the in reason behind the success of this algorithm.

Now, here so, so idea it is there. So, since the total population of size 2 N, and we have obtained the many fronts, then all fronts may be may not be accommodated in the fronts available in P dash. So, all fronts which can could not be accommodated are simply rejected. Now here again I can say this one say F 1 F 2 dot dot dot Fk dot dot dot F m. So, these are the all fronts that can be obtained from R so, given R non-dominated sorting procedure will calculate all front.

Now, up to which the Fk front can be filled so, to the current population P dash like ok. And so, that it is less than the size of P that is the population size now. So, the remaining front those are there they can be simply ignored so, they simply reject it. So, this is the procedure it is there, because they are not good for creating the mating pool or next population generation.

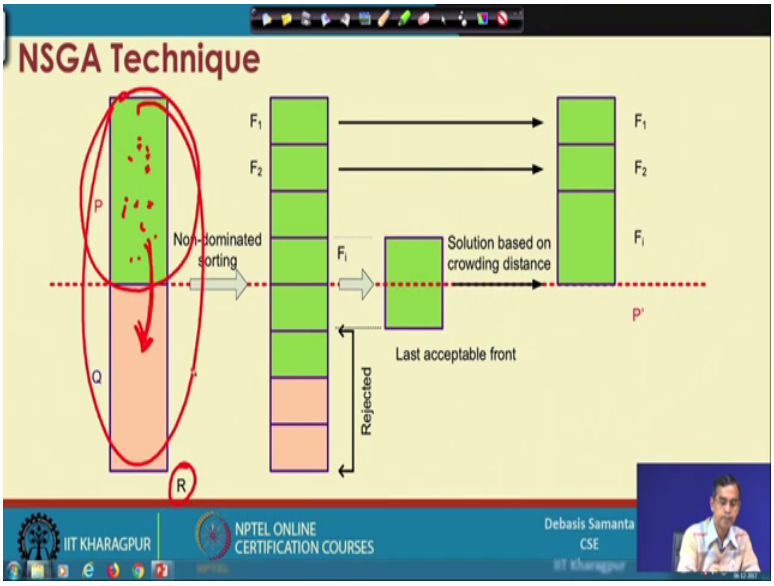
Now, here the idea again, instead of arbitrarily discarding some members from the last front, the solutions which will make the diversity of the selected solutions are the highest R should be chosen. So, this is the one important criteria or rationale that so, the last front contains some solutions who is basically then the maximum size of the population, then how to select the solution so, that it exactly the same as the population size. So,

basic idea is that the from the last front, we have to select thus that solution which has the very good so far, the population diversity is concept.

Now, in order to select those solutions, which is there in the last front the NSGA 2 procedure, consider one method it is called the crowding distance method. And this crowding distance is basically is idea about that how to select a winner out of the solutions though is those are there in the last front. And so, so this is the basic idea about here, in this technique the crowding distance we will discuss in details in the next slide. So, crowding distance is very important concept there we will discuss these things in the next slides.

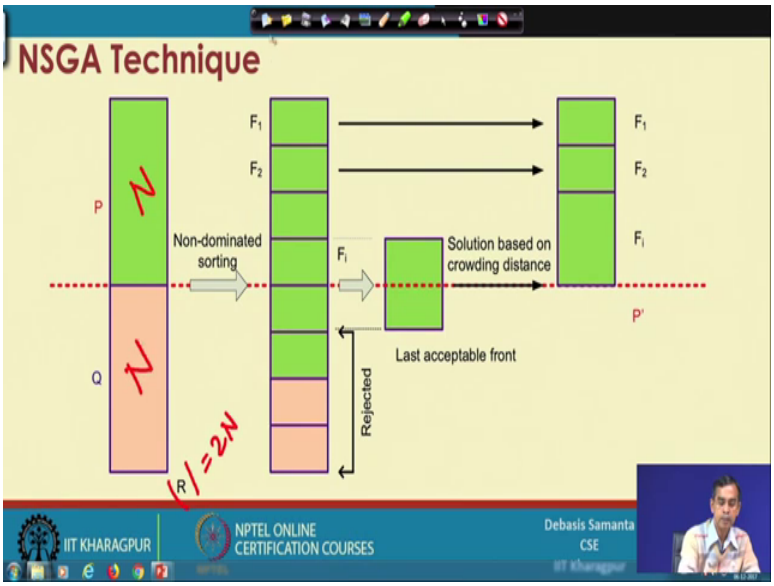
Now, let us first proceed about the other concept it is there ok.

(Refer Slide Time: 16:47)



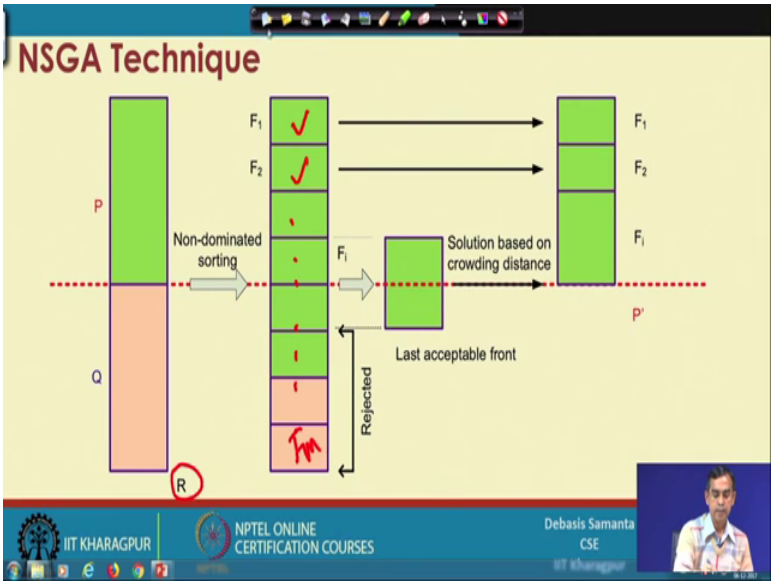
We can summarize the discussion about the basic steps that is there in NSGA 2. So, it is basically NSGA 2 technique, right? Now here we can summarize the thing that we have discussed here. So, these are the current solution in the current generation. So, P and from the P we produce the Q the next, I mean solutions for the next generation. So, P and use the reproduction method so that from P the Q can be generated. So, all the solutions together it is called R and then the size of R equals to 2 to the 2 N, where N is the size of P and N is the size of Q.

(Refer Slide Time: 17:36)



So, these basically the solution sets, that needs to be considered for the NSGA 2 approaches. So, these are first step; that means, the current solution P, and then reproduction after reproduce solution Q are combined together. Now from this current solution P and Q or that are the mark solution R. We have to find the non-dominated front from here.

(Refer Slide Time: 18:10)

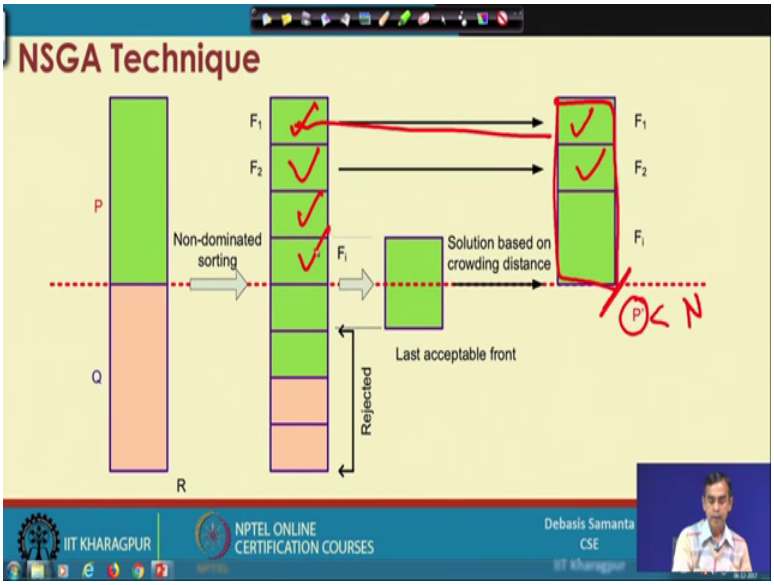


So, so suppose these are the front F 1 F 2 dot, dot, dot, dot say F m number of fronts are there. So, they are the separate fronts has been calculated ; that means, this is the front

one first front 2 second front 3 third and so on. So, this is this can be obtained following the non-dominated sorting procedure.

Now, next is basically selection. Now this selection first is that this F 1 front should be selected if we see that after adding this one, the size of the population will not exist the quota or limit that is the N.

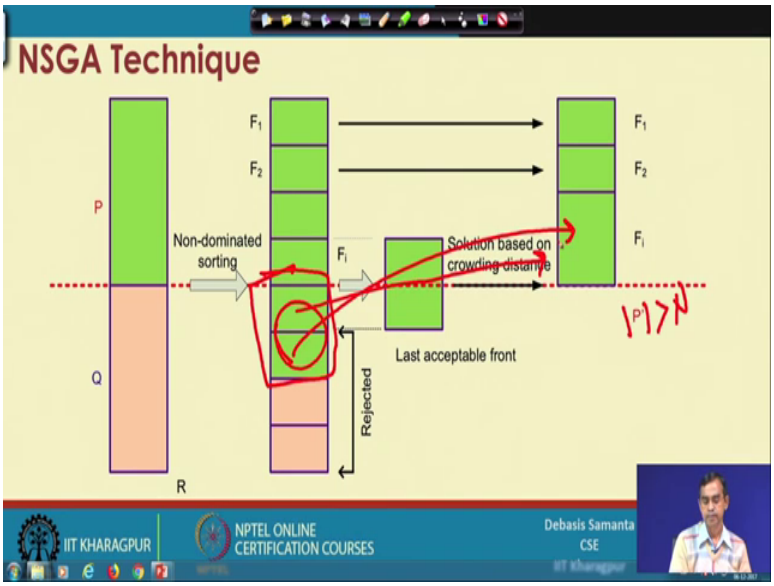
(Refer Slide Time: 18:40)



So, it is basically P dash so, we have to go on. So, if F 1 will be added into this P dash if we see that size of the P dash is less than N. So, the if this is selected similarly F 2 can be added here, if we see that the size of the P dash is less than N. So, this way we have to select all this one.

Now, after selecting so many things suppose at this space.

(Refer Slide Time: 19:17)

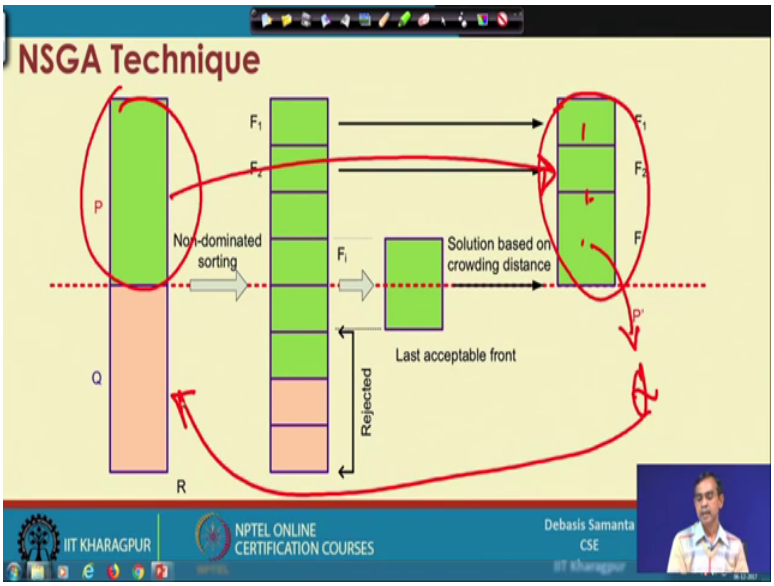


One front is like this, which is the next one to be considered, but if we add this front here, then the size of the P dash will be greater than greater than N. This means, all the solutions which is there cannot be arbitrarily selected into here, right?

So, this is the idea about then this procedure says that then we have to play one tournament among all the solutions which belongs to there, and based on the tournament selection we will select the solutions here to fill the population size P dash is equals to N.

Now, this tournament selection is basically based on one technique it is called the crowding distance technique. And all the solutions so, from there all the solutions which are the winner to be selected here, and for the rest of the solutions are simply to be rejected.

(Refer Slide Time: 20:23)

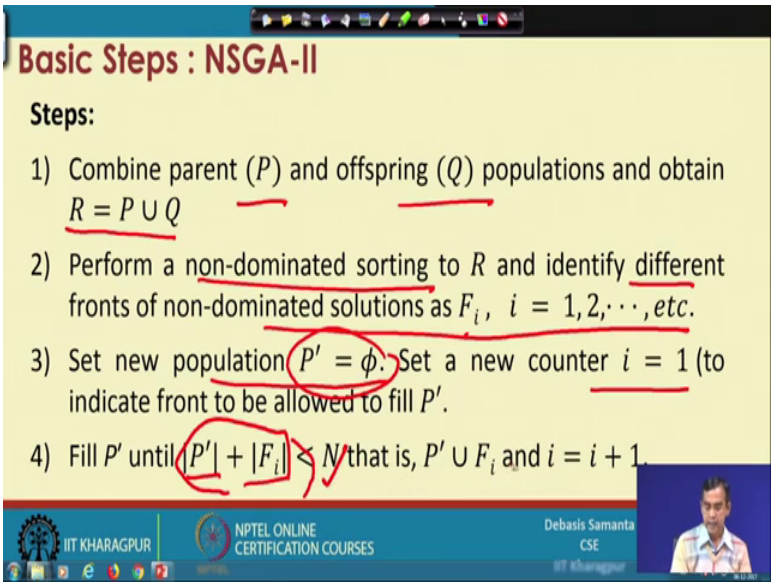


So, this way from the current population P, we will be able to obtain the selected population, and this is basically the mating pool that needs to be considered for the next generation population.

So, from this mating pool we will be able to consider the queue, the next generation population and then same procedure will be repeated again and again until the termination criteria is satisfied. So, this is the idea that the NSGA techniques follows here, but in order to understand the NSGA technique again ah; that means, here is this procedure is the most important procedure; that means, how the tournament selection based on the crowding distance can be obtained.

So, this is the important thing that is there.

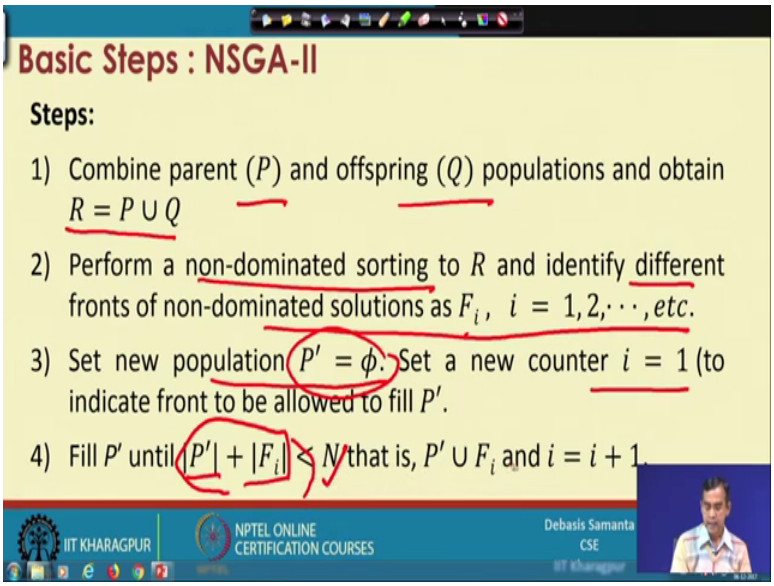
(Refer Slide Time: 21:14)



And so, basic idea again so, whatever the step that we have discussed earlier it can be little bit expressed in a formal way combined parent P, and then offspring solution cube to produce the resultant solution R, perform a non-dominated sorting to R, and identify the different front this is called the non-dominated front F I like 1 to 2 etcetera.

Now, set new population P dash initial which is empty. And then we have to follow this one, until we have to select the front. Now P dash plus front one should be alert. So, that it is less than N and we stop this selection or adding, when we say that it greater than N. So, for the greater than N, we have to consider the method; that means, say selection method that is called the based on the crowding selection.

(Refer Slide Time: 22:04)



It is called the crowded tournament selection procedure.

Now, the crowed tournament selection procedure follows one method it is called the crowding sort. For all the solutions which belongs to Fi. And then based on this crowding sort, it will consider N minus P dash solution, because these are the solution that needs to be filled up, and that can be based on the crowed tournament selection. And this way the next generation can be obtained and then next offspring generation can be obtained.

So, this is idea about it, now learning the NSGA 2 for this method, that mean how to select the solutions from the last dominated front in order to make the solution size same as the population size. And then this procedure needs the discussion of the crowding sort techniques. So, the crowding sort techniques so, we will discuss about the crowding sort techniques in the ok.

So, we will discuss the crowding sort techniques in the next lectures.

Thank you.