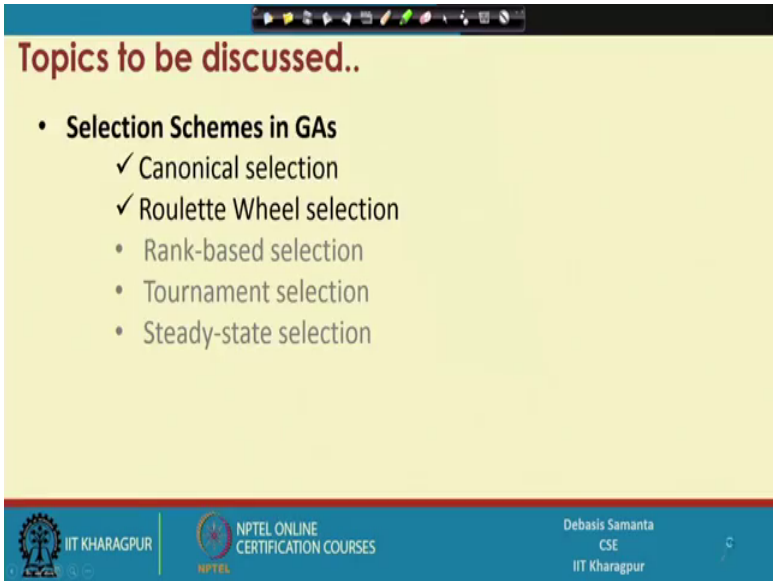
Introduction to Soft Computing
Prof. Debasis Samanta
Department of Computer Science & Engineering
Indian Institute of Technology, Kharagpur

Lecture - 19
GA Operator : Selection (Contd.)

We are discussing the selection operation in genetic algorithm today we will discuss the few more applications, few more selection operations. So, the in the last lecture we have discussed about two selection operations, namely canonical selection and roulette wheel selection, we have learned there are certain limitation in roulette wheel selection.

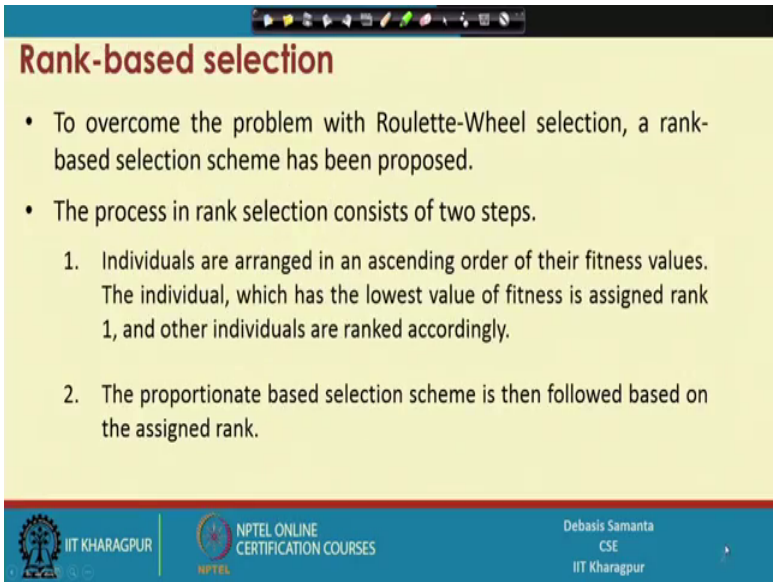
The other selection operations that we are going to discuss it today, these are rank based selection the tournament selection and steady state selection.

(Refer Slide Time: 00:56)



So, let us first discuss about rank based selection.

(Refer Slide Time: 00:56)



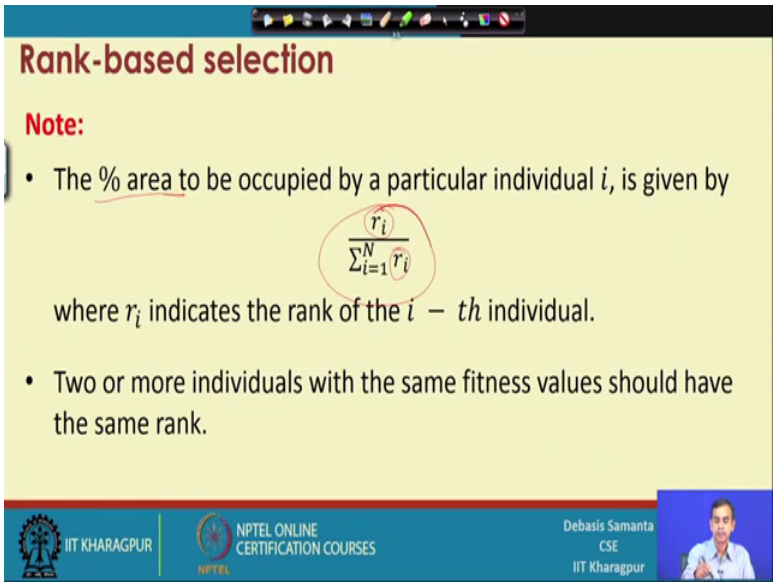
The rank based selections have been proposed to adjust the problem that is there in the roulette wheel selection. We know that in the roulette wheel selection basically the individuals, which has the highest fitness value is preferred more compared to the lowest fitness value.

Sometime this is not a fair selection rather we should give a chance to select the inferior population also, because sometimes the offspring that will be produced by the mating of the best solution and then worst solution can lead to a better solution with faster termination. Anyway, so rank selection is basically best the basically proposed to remove the biasness towards the high fitness value populations. So, in this process these processes consist of brought two steps.

So, in this process all the individuals are arranged in an ascending order of their fitness values, and then the individual which has the lowest value of the fitness is assigned as rank 1 and then the next lowest fitness value is rank 2 and so on.

Once the rank is assigned to each individual then we follow the any proportion of the selection scheme likes a canonical or roulette wheel then, and this way we will see how the favour towards the highest fitness value that is given by the roulette wheel have been checked in rank based selection method.

(Refer Slide Time: 02:50)

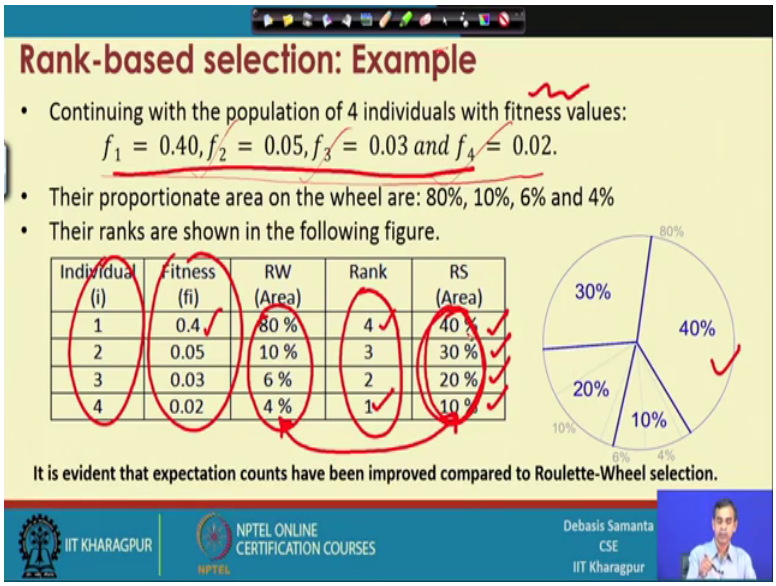


Now, so the rank that will be assigned it basically decides the percentage area to be occupied by a particular individual say it is i and the formula that we follow to for this purpose is it is written here.

So, it is the same as the fitness value or fitness core calculation that we have learned in roulette wheel scheme, there we basically r i was replaced by fi there and so there. And this basically denoted as a pi in the roulette wheel selection strategy, but here it is the calculation of the rank and based on this rank, we decide the probability that it will be selected for the things. Now, here the percentage area basically indicates as in the roulette wheel; that means, it is the proportionate that it will be I mean it is eligible to select for the next generation.

Now here there may be sometimes that i; that means, two or more individuals with the same fitness value is the quite possible, in that case we should assign the same rank to the individuals ok.

(Refer Slide Time: 04:20)



So, after assigning the rank our next task is to up is to follow, either roulette wheel method or canonical selection method, but not based on the pi or fitness there rather it is based on these ri; that means, the rank values now let us illustrate the rank based selection scheme with an example and in. So, let us consider only the 4 individuals at any instant, and they are f 1 f 2 f 3 and f 4 and the value is shown here, f 1 f 2 f 3 and f 4 then ok.

Let us illustrate the rank based selection with an example, in this example we consider 4 individuals, here as we have shown the 4 individuals here f 1 to f 4 with their different fitness values, and the same populations are listed here the 4 individuals and their fitness values listed there. And so these are the rank that we have assigned so the individual with the lowest fitness is here. So, the rank is 1 the highest here the rank is 4 and so on, so on.

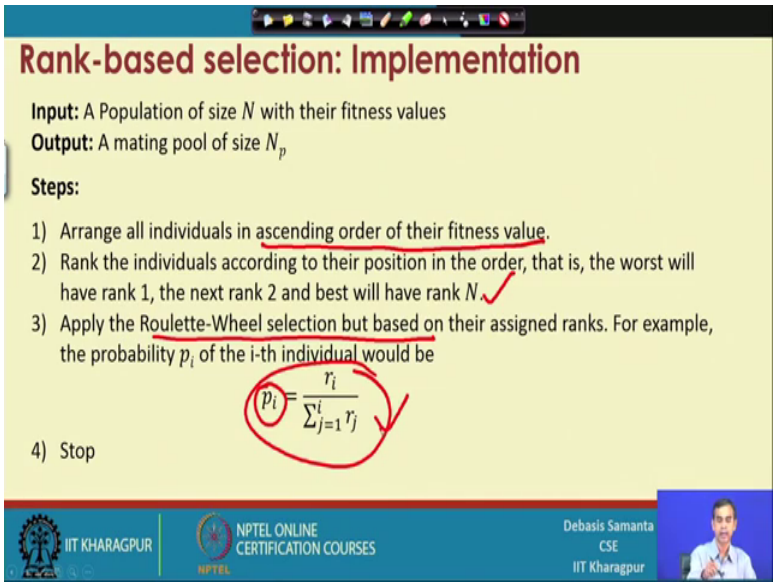
So, this is the rank assigned for the each individual and then on the basis of rank and this is basically the calculation according the rank selection method about their score. So, it is basically point 4, point 3, point 2, point 1, is basically rank divided by some of all the rank like so these are the things are there.

Now, we will apply the proportionate based any selection say roulette wheel on these values there. So, typically it is basically like, look like this one, now if we follow the same thing, but only without any rank then roulette wheel selection will assume these are the values, so the difference is in there right.

So, roulette wheel selection follow these are the course and whereas, the rank selection follows these are the course. So, we can see the 80%, which is considered in the roulette wheel it become consider a 40% at the score and, so on, so on. So, the score has been little bit changed because of this rank calculation and then the rank based selection strategy.

So, this way we can give some favours to the worst individual event for example, earlier which was has given the weightage 4%, it is basically 10% and then also favour towards the highly feet values also reduce 80% to 40%. So, this is the mechanism here it is followed, and if we follow this one the fair chance that the inferior quality, inferior individuals also will be selected compared to the superior in individuals as it is always there in the roulette wheel scheme.

(Refer Slide Time: 07:23)

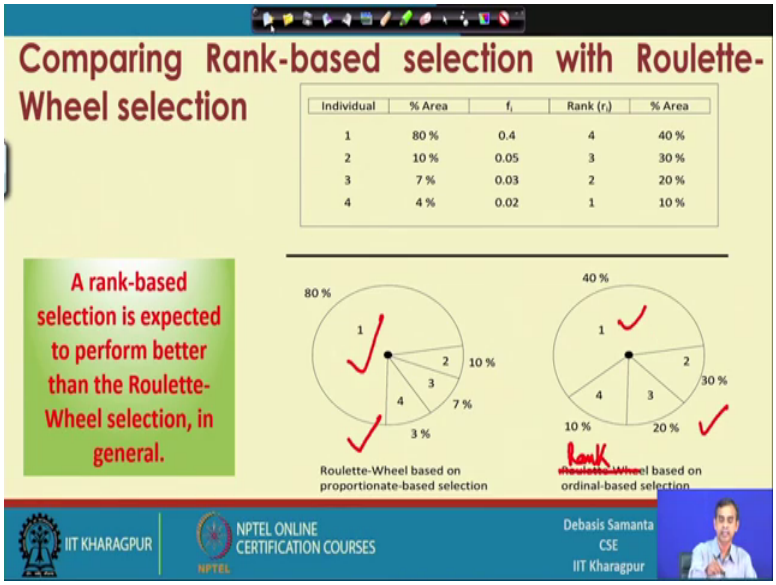


So, this is the basic concept of the roulette wheel scheme. And so for the implementation of this algorithm is concerned we can explain it very briefly. So, the first step is as we have learned about the first step is arrange all the individuals in ascending order of their fitness values, then rank the individuals according to the position in the order that is the worst will have the rank 1 and the next will rank 2 and then best will rank n so on.

Then once the rank is assigned we have to follow the roulette wheel selection, but based on their assigned rank value which will be calculated like this one. So, basically pi you

can note it the pi was calculated in roulette in a different manner, but here it is calculated based on their rank, so this is the procedure that we can follow.

(Refer Slide Time: 08:16)

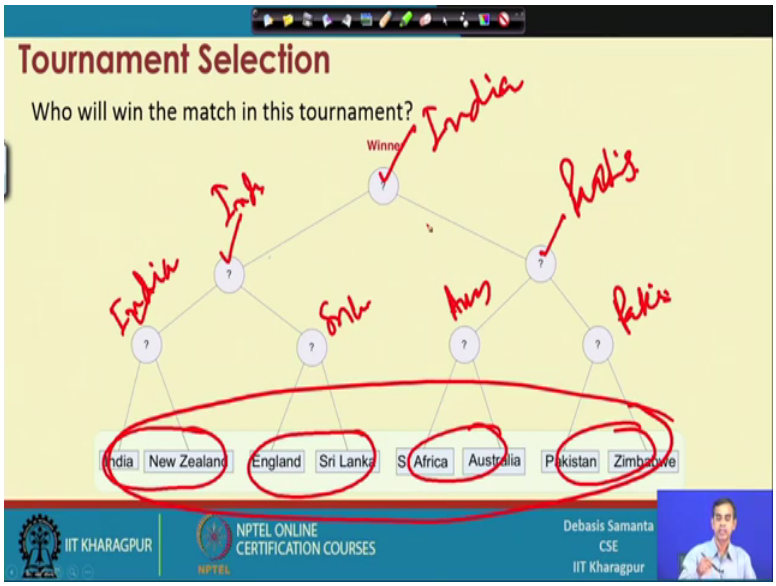


Now, let us see how it is beneficial compared to the wheel, roulette wheel based selection. So, the idea it is basically if we follow the roulette wheel selection the wheel game is like this and if you follow the rank based selection the wheel game is like this so it is basically rank based selection ok.

So, in this, so we can say that for the roulette wheel the fitness value having 80%, has the most probable that it will be selected whereas, the probability is reduced here as the area is reduced here.

So, this is why in general we can say that any rank based selection is expected to perform better compared to the roulette wheel based selection in general it is the result that can be obtained, and then it is proved that usually the rank based selection is better selection strategy than the roulette wheel selection.

(Refer Slide Time: 09:15)



Next we will discuss about another selection strategy it is totally different than the selection strategy that we have discussed those are the basically based on the proportional based algorithm, here we have we are going to discuss a new strategy based on the tournament strategy.

Now, so we know the tournament strategy, here there are different type of tournament that can be played, I have given an example here it is basically the knockout tournament. So, in case of knockout tournament the idea it is like this the different games are planned for example, here is the games between the two, what is called the teams is a another, another and different games are there.

Now, so many games out of the different what is called the players are selected then. So, a game will be played for between the two players here. So, India and New Zealand and who is the fittest here will be selected here, so say India is the fittest India is selected.

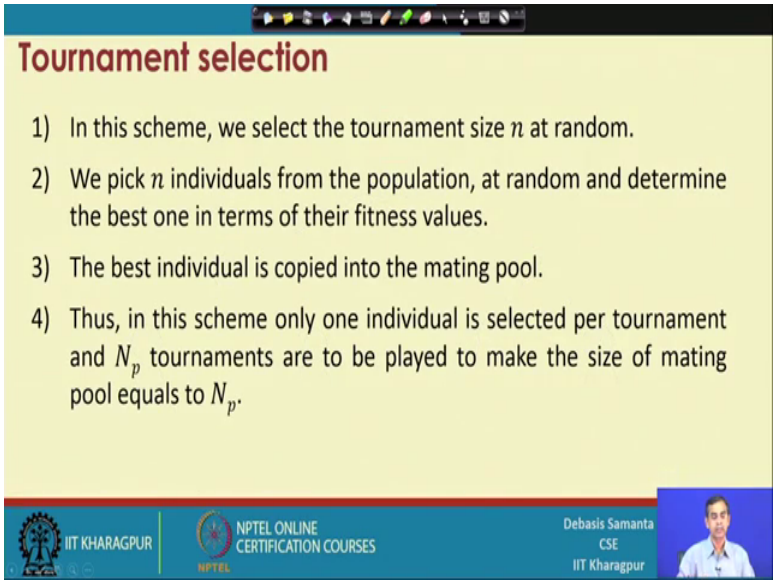
Similarly, another play will be played here and Sri Lanka and England. So, Sri Lanka is selected as a Sri Lanka is most fitted suppose compared to the England then the another play. So, this way similarly here also Australia and here is the Pakistan is selected because of some fitness value there.

Now, again another game is played, and so the here India versus Sri Lanka suppose India is the feast. So, in this game India win this and here in between Australia and Pakistan.

So, in these game say Pakistan is more fittest than the Australia. So, Pakistan is selected and finally, the play between India and Pakistan. So, here the winner say India is more fittest than Pakistan, so India is selected.

So, this way if we play the game among the different players here and then after the end of the tournament the best, I mean the best the player with the best fitness values will be selected. So, this is the general procedure of the tournament selection strategy, the similar strategy has been followed in our GA, operation also we will discuss this tournament strategy in GA.

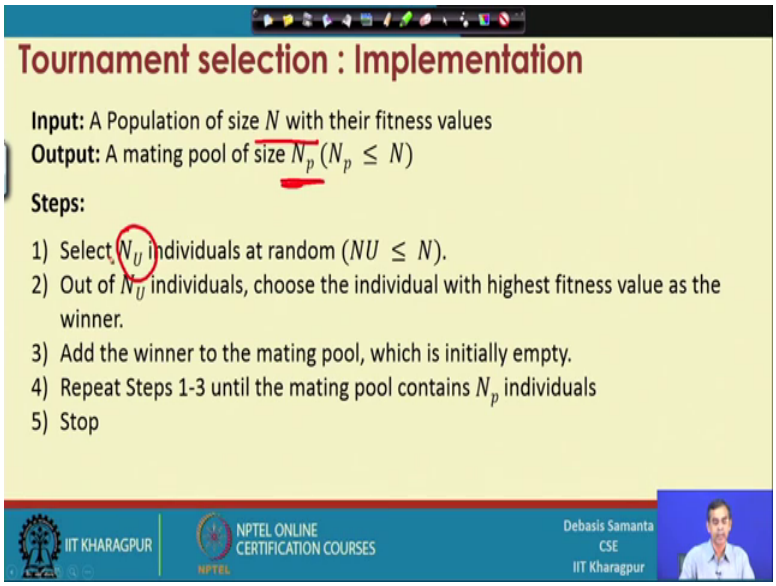
(Refer Slide Time: 11:27)



And that is called the tournament selection procedure. Now, so the GA tournament selection procedure has few steps here we have shown the 4 steps in this scheme we select arbiter is the team size N at random; that means, say suppose population size is 100, let us decide the team size is 10 say then out of this 10 team.

We have to play the tournament as we have discussed earlier knockout like and then we have to select the 1 individual and then if we repeat this kind of tournaments for Np times, where Np number of population to be selected for the mating pool then we will select this one. So, the idea it is like this.

(Refer Slide Time: 12:09)

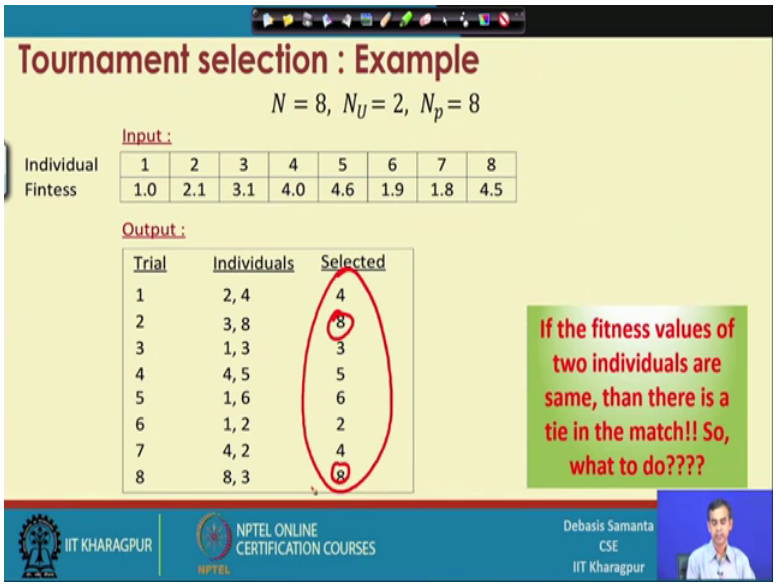


Now, so this is the idea and that algorithm to this idea can be stated like this. So, input is a population of size N, and their fitness values are calculated for each individual and the output should be a mating pool of size Np, where Np is some values less than N.

So, according the strategy we have to select nu individuals at random. So, you can select this nu number individual at random and nu should be very large, very small value compared to l N. The total size of the population once the any individual is selected out of any individuals we have to play the tournament knockout tournament like, and then the individuals with the best fitness values will be selected as the winner.

So, the winner which has been selected will be added to the meeting pool which is initially empty. Then we will repeat this step 1, 10, 3, for Np times, until Np individuals are selected for the mating. So, this is the straightforward procedure, in fact, only the calling procedure here in the step 2 which basically plays a tournament among the nu individuals, but this is also not, so costly and it is manageable and affordable.

(Refer Slide Time: 13:36)



So, this is the I mean strategy that we have discussed let us illustrate this strategy with an example and. So, here in this example we can see n the total 8 number of individuals are there and we have to select say, 8 individuals for the mating pool; that means, Np equals to 8 and n is also 8 here.

Now, we select nu size of the tournament is 2 for simplicity and then illustration purpose. So, here the first trail; that means, we select first n individuals from this 1 at random. So, selected 2 and 4 is selected and out of these 2 two and 4 the winner is 4 because it has the highest fitness value compared to these one.

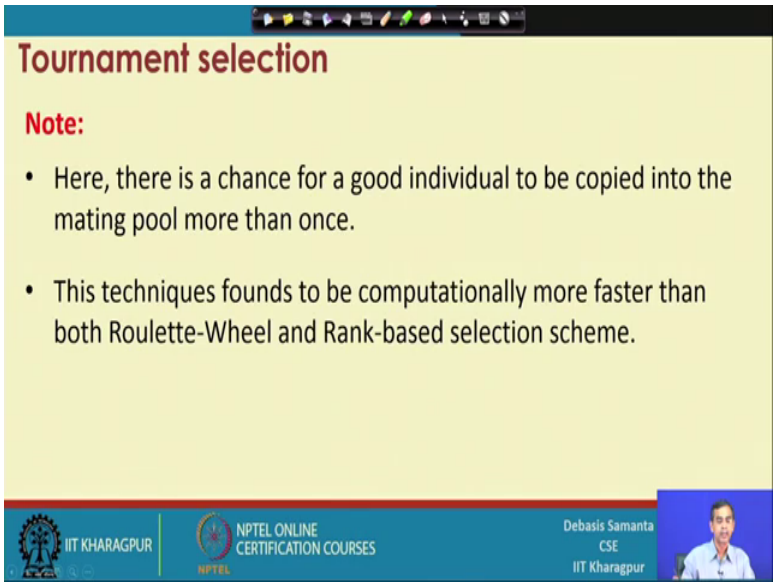
So, 4 is selected, so this is the first trail similarly second trail 3 and 8 is selected and out of these 3 1 4 point 5 having the high so 8 is selected. So, this procedure is continued and these are the 8 difference select say population that can be selected for the mating pool, here we have intentionally used N and Np are the same values.

So, here we can see that according to this procedure, the same individual may be selected more than once for example, here the individual 8 is selected here and here more than once. So, it is quite possible that it can be selected the same things more than once.

Now, again here for example, there may be another twist another what is called the tie also suppose 2 individual selected although all individuals are having the same value say suppose 6 and 7, suppose any 2 individual say 4 and 5, it has the value is 4.0 right, and in

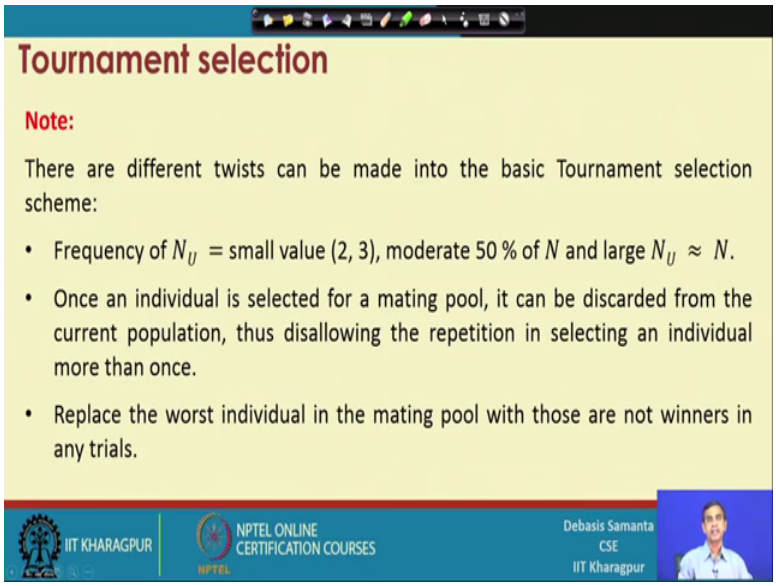
a trial if we select the individual 4 and 5 for the tournament for example, here 4 and 5 in the tournament then both have the same value then which I mean individual needs to be selected, so whether 4 or 5. So, in case of tie we can break the tie of certain tossing mechanical and then anyone at random can be selected ok.

(Refer Slide Time: 15:51)



So, this is the tournament selection strategy, and in the tournament selection strategy what we have learned is that, there is a chance for a good individual to be selected into the mating pool more than once that we have already mentioned and this technique founds to be computationally more faster than both roulette, wheel and rank based selection scheme and it also has many other benefits or the benefits of the different selection strategy and their limitations will be discussed at the end of the lectures today, so this is the tournament selection strategy.

(Refer Slide Time: 16:25)

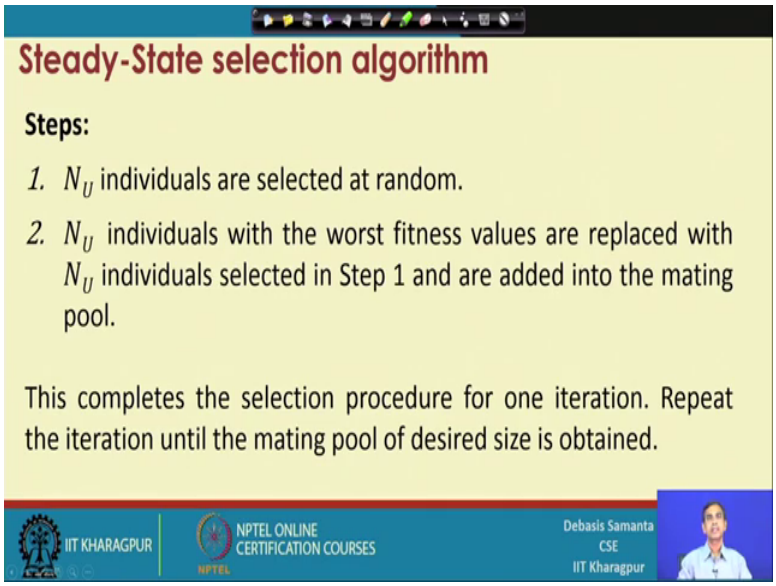


And tournament selection strategy also has certain twist we can follow to make it more applicable and appealing. So, there are few twist which have been hinted here, so the nu can be changed any value as small as 2 and then as large as N. So, so accordingly the different results can be obtained and then the programmer has to choose which results is favours to give the best result, and then at the fastest execution of the genetic algorithm.

And then another twist can be, so in that case it is repetition is allowed we can check this repetition, this means once an individual is selected for a mating pool it can be discarded from the current population thus disallowing the repetition in selecting the individual for more than once.

So, that is also possible if we want that only 1 individual will be selected only once and another twist also can be added there in this strategy, if we replace the worst individual in the meeting pool with those are not winners in any trail. So, the worst individual can be replaced by some individuals in the mating pool which are not winners in the trail. So, these are different twist can be followed to make the tournaments selection more robust and more reliable in a different GA execution.

(Refer Slide Time: 17:44)



Next we are going to discuss the another strategy, it is called steady state selection algorithm it is a very simple most strategy it works sometimes very effectively, but not always, but it has computability is very fast and then result is also more or less acceptable compared to the other strategy. So, some programmer follows this kind of strategy also.

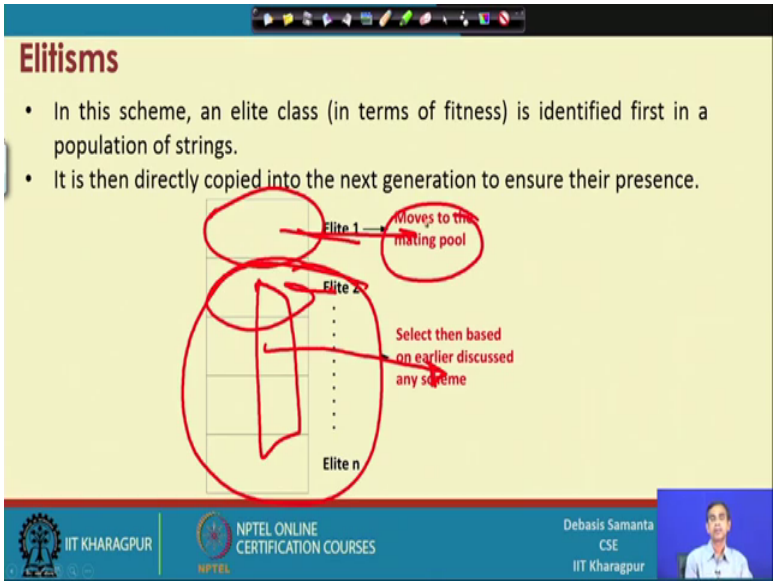
Now, this strategy is basically like tournament selection it also selects 1 parameter call nu that is the numbers that can be selected. So, nu can be as small as 2 usually it is very small number compared to the number of size of the population. So, nu individuals are selected and they are selected at random. So, randomly any 2 individuals are selected from the current population.

Nu individuals with a worst fitness values are replaced with the newly selected any individuals in the previous state and added to the mating pool. So, this procedure is repeated for Np times, where Np is the number of mating I mean populations to be selected for the mating. So, it is the procedure is like this and if you see as the number of iterations is increase it always refine the current population by replacing the worst.

So, worst will be I from there and then all the based population will be selected and that is giving more fair chances to the individuals with higher fitness value it is selected here compared to the worst one.

So, this is the procedure it is there, it is the simple most procedure one, but it is not so much what is called the control.

(Refer Slide Time: 19:28)



As it is possible there in other selection strategy, but as a simple most strategy it can be followed in some application now. So, our objective was is basically to select the individuals which can be played for the mating pool, and we have discussed the 4 different strategies, the canonical, roulette, wheel and then rank based selection. whether the population based and then finally, to selection of the tournament and steady state. So, 5 selection strategies that we have discussed it.

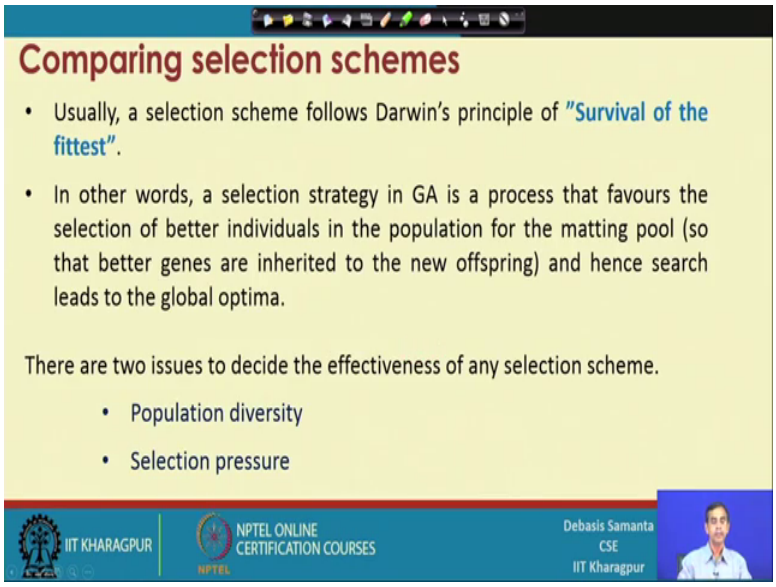
Now, So, there are again few what is called modifications in the selection mechanism is followed in recent gas, and this 1 such mechanism is called elitism what is the meaning idea is that be depending on the fitness values of the individuals the individuals are grouped into a number of elite group, so elite 1, elite 2, elite n, so on, so on. So, here basically the idea is that a set of individual based on their fitness value belong to the elite 1 then other is elite 2 and, so on, so on.

In other words the individuals which belong to the elite 1, they are highly fit what is called population or highly fit individuals. So, the strategy according the elitism that we have to move all the individuals which belongs to elite 1 to the mating pool, then the then we will select the remaining in order to make the size of the mating pool as Np. So,

select them from the remaining elite groups, but following some strategy either the roulette wheel or rank based selection or tournament selection there.

So, this will be selected whatever the existing strategy to select the waste of the things and then passed without any selection they are moved to the mating pool. So, this is the 1 strategy that is followed and it observe that this strategy also very much effective in some situation.

(Refer Slide Time: 21:35)

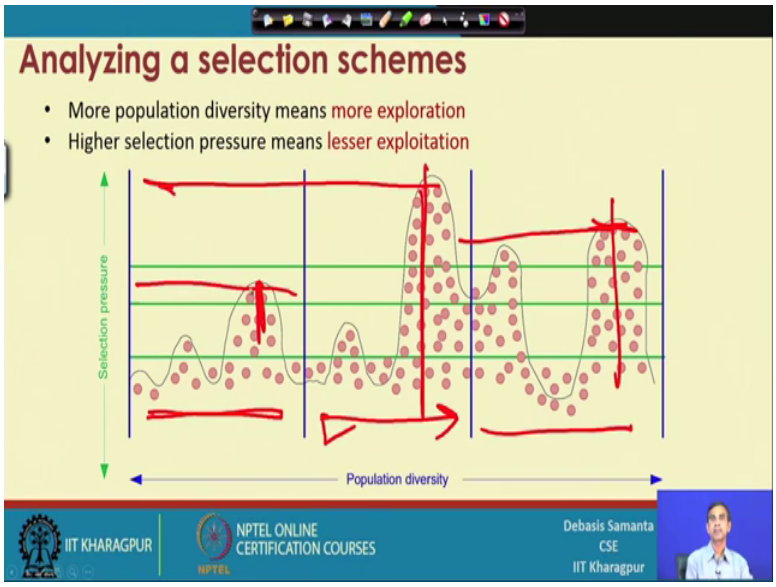


Now, in in short we will compare the difference selections scheme. So, basically the selection scheme follows the Darwins principle of survival of the fittest; that means, the individuals which are the highest fitness values should survive better compared to the worst fitness best individual.

So, this is why the selection is followed there and all the whatever the selection strategy basically targets; that means, that so that we can follow the Darwins principle there. And as we have learned it in a selection strategy, so all selection strategy anyway more or less favours the selection of better individuals in a current population and from where the mating pool can be selected; that means, it will allow mating to those individuals they are basically the best individuals. So, assuming that from the best individuals another best offspring will be obtained.

Now, in this regard the any selection scheme rather can be I mean compared or their effectiveness can be compared in terms of 2 concept 1 is called the population diversity and then selection pressure. So, in order to compare the different selection scheme that we have learnt first we have to understand about what is the population diversity, and then selection pressure. Now, let us understand the first population diversity and then then we will discuss about selection pressure.

(Refer Slide Time: 23:10)



Now, population diversity means if the ranges of fitness values are within a wider range, then we say that the population has very wide diversity.

So, as an example in this figure, so the range of fitness values is shown in this direction. So, these basically shows the population diversity or in other words if 1 population is in 1 population is this 1 only then it has lesser population diversity than the entire one. So, population diversity indicates that how many different that how what is the wider range of the fitness value that the population currently having.

Now, population diversity implies more exploration; that means, we can get the more accurate result for example, if we consider this is the 1 population right and if we ignore this 1, then may be that GA will start at with these optima, so this is the optimum one.

However if you consider the population this 1, then we may start at this 1 also. So, if we don't have the wider population diversity then we may trap into local optima. So, in

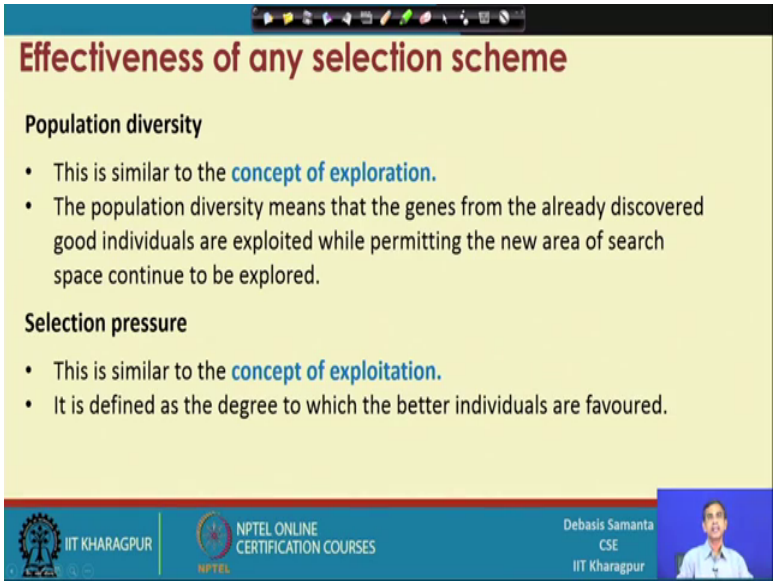
other words we can say the more population diversity means more exploration now next is selection pressure.

So, the selection pressure is basically, so it is basically what is the highest values that is there in in a current population for example, if we consider, so this is the 1 population then selection pressure is this 1 this basically because highest fitness value this 1. On the other hand if we consider this is the population then this is this is the selection pressure, selection pressure is this 1 and this population has a selection then highest values of signal will this one.

So, here three different populations that we have consider having the three different selection pressures. So, here basically selection pressure means a lesser exploitation because whenever we get the selection pressure for example, if we take this population and selection pressure is this 1 then our optimum value will be confined into these kind of things one.

So, basically will not be able to exploit the better results that can be had from the from the if the selection pressure is something else. So, this is the concept it is here, so in summary we can say that more population diversity means more exploitation and higher selection pressure means it is basically the lesser exploitation. So, it is basically exploration versus exploitation in the mechanism.

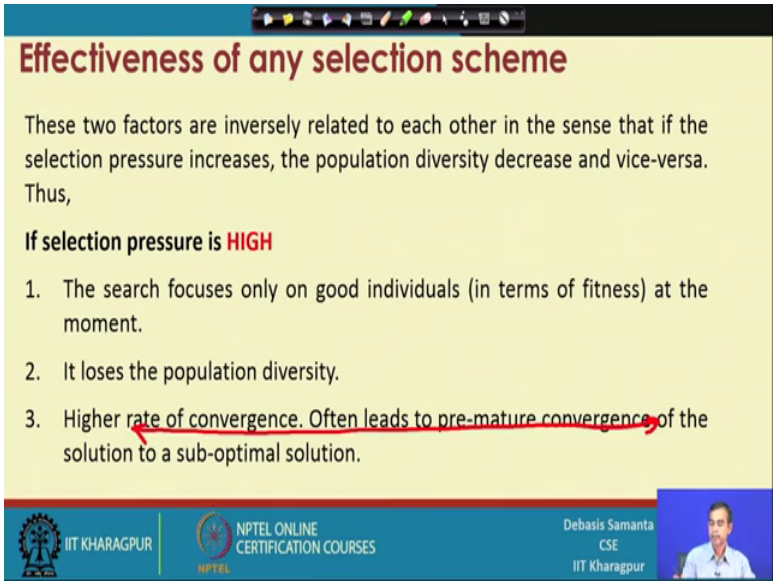
(Refer Slide Time: 26:00)



Now, let us see the how we can compared the different selection scheme anyway. So, population diversity as we have learnt this is similar to the concept of exploration as we have already mentioned, and then selection pressure is basically concept of exploitation, this means that it is defined as the degree to which better individuals are favoured.

On the other hand, so for the population diversity is concerned it means that genes from the already discovered good individuals are exploited while permitting the new area of search. So, basically the idea it is there if we follow the wider population then we have better searching capability, because we can search this one also we can search this one also and. So, many direction we can search it right, anyway so these are the concept it is there.

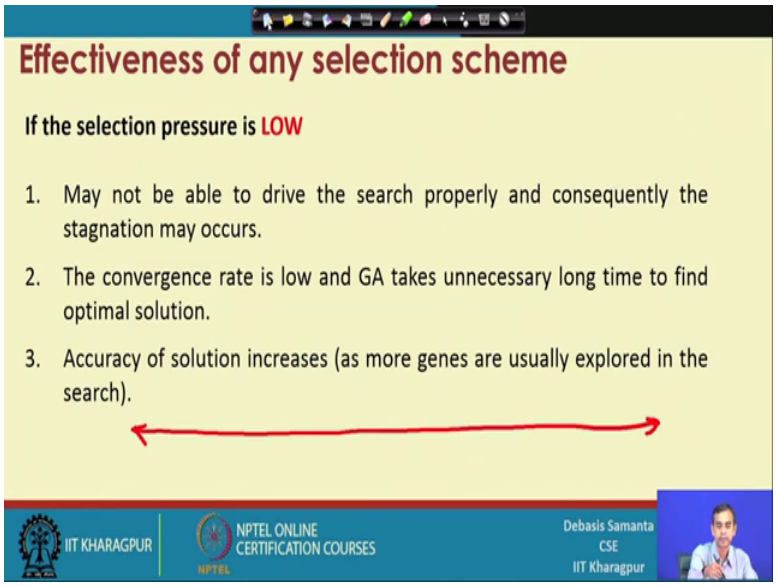
(Refer Slide Time: 27:00)



Now, we will quickly mention few points about if the selection pressure is high or low. So, first let see if the selection pressure is high then the search focus only on good individuals in terms of the fitness values at that moment and it therefore, loses the population diversity and then it basically leads to the higher rate of convergence that mean it will converge at the fastest rate and whenever the convergence is fast it may be premature convergence of the solution and therefore, the solution may not be, so accurate.

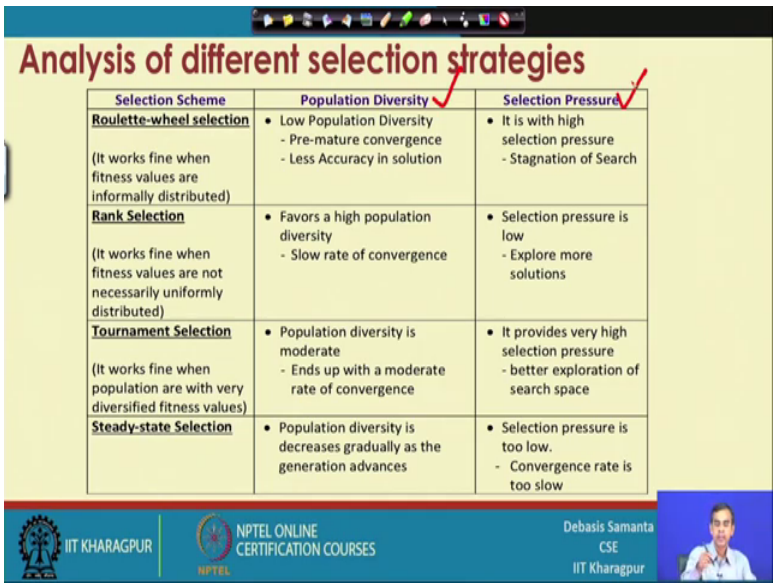
So, there is a basically trade off between the selection pressure and the population diversity here.

(Refer Slide Time: 27:51)



And on the other hand if the selection pressure is low, so it may not be able to drive the search properly and consequently the stagnation may occur and the convergent rate compared to the selection pressure is high is low and GA takes unnecessary long time to find optimum solution, accuracy of solution; however, if the selection pressure is low increases as more number of individuals are explored in the search process.

(Refer Slide Time: 21:20)



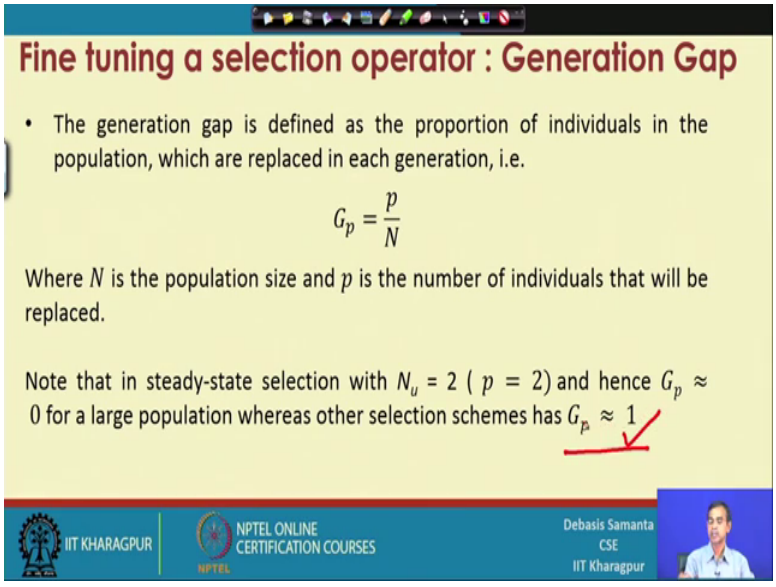
Now, I will just this is the last slide and I will quickly summarise the different what is called the techniques there in terms of population diversity and the selection pressure

now as that we will says that. So, for the roulette wheel selection is concern it provides low population diversity, for selection pressure is high.

On the other hand, in case of rank selection it favours high population diversity and then selection pressure is low. And if we consider tournament selection popular population diversity is moderate and it provides very high selection pressure. Steady state selection on the other hand, population diversity is decreases gradually as the number of generation in advances where the selection pressure is too low.

So, this comparison gives us enough idea about which selection strategy are to be followed, if we want that high population diversity then definitely we can go for tournament selection on the other hand, if we want very high selection pressure and then population diversity low then we can I mean choose for roulette wheel selection strategy. So, these are different selection strategy, that is there in the GA theory we have discussed in this lecture.

(Refer Slide Time: 29:39)

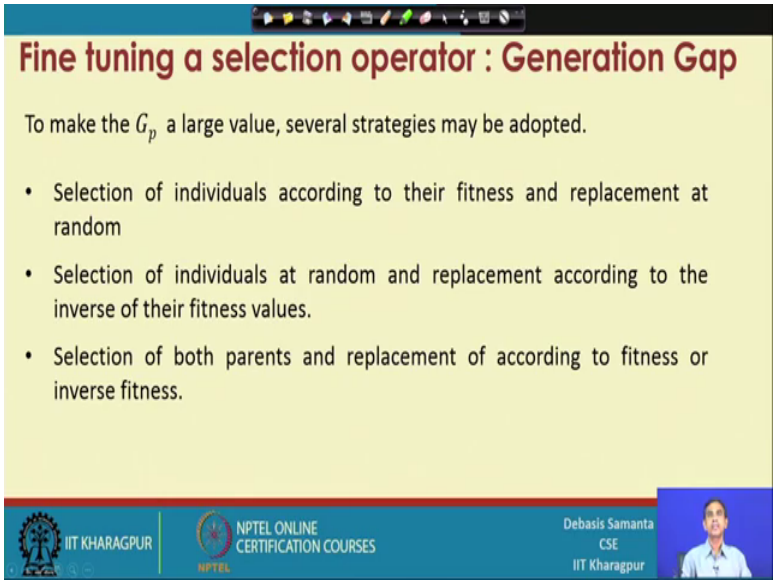


And we will discuss about other operators also there, it is basically called the generation gap, generation gap is denoted by p by N. Where p is basically the number of individuals those are basically will be replaced and N is basically the number of population is the size of the population.

Now, for example, in case of steady state selection, so Gp equals almost 0; that means, because if we take n equals to 2 and for a large population size, on the other hand for other selection strategy Gp equals to very high it is as high as 1. So, that the maximum value and the lowest value is 0.

So, Gp lies between 0 and 1, and Gp also can be considered as a measure about a selection strategy and it is usually preferable that selection strategy, which has better generation gap. So, in that case a steady state selection is not preferable because generation gap is very less ok.

(Refer Slide Time: 30:40)



We have consider the different selection strategies here, and our next topics that we are going to discuss is another operation, another GA operator it is called crossover operation.

Thank you.