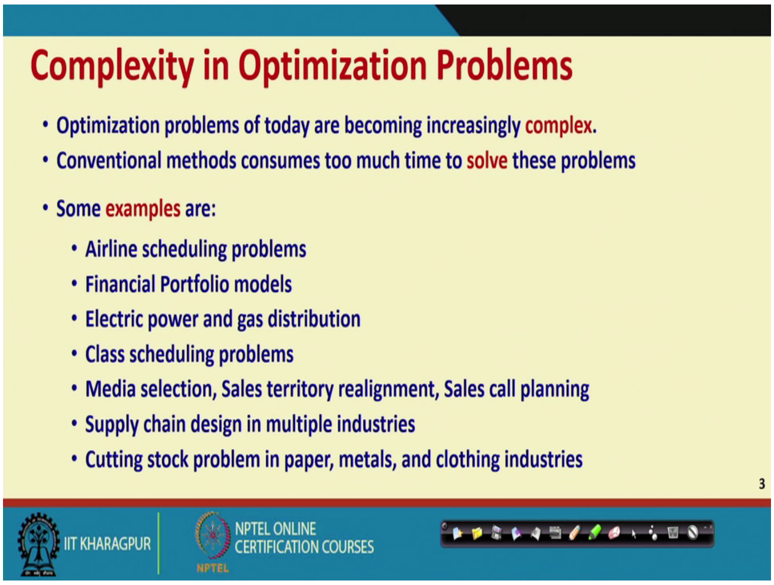
Selected Topics in Decision Modeling
Prof. Biswajit Mahanty
Department of Industrial and Systems Engineering
Indian Institute of Technology, Kharagpur

Lecture – 31
Introduction to Metaheuristics

Right so, in our course; Selected Topics in decision modeling, we are in our thirty first lecture and now we are going to start a new topic on Metaheuristics; right. So, as a part of introduction, let us have the introductory part on the Metaheuristics.

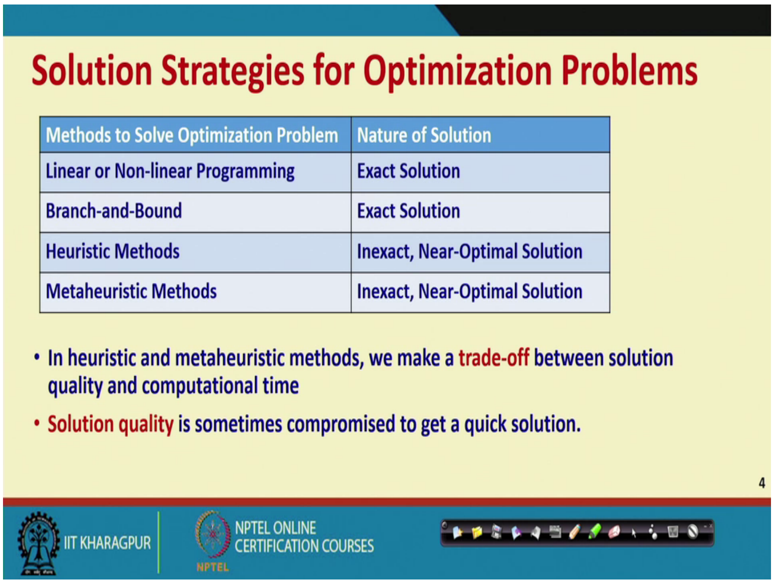
(Refer Slide Time: 00:40)



Now, so, the essential idea is the is that we have large number of optimization problems that we have to solve in different fields and increasingly, this problems are becoming complex is, alright. So, what happens any attempt to solve them by conventional methods? Either fails to really solve the problems or they become too time consuming. So, when I say time consuming I mean probably even in years. So, here are some examples.

Examples in let us say in airline scheduling problems, financial portfolio models, electric power and gas distributions, then scheduling problems of different type media selection, sales territory realignment, sales call planning, supply chains or in cutting stock problems where you have to cut you know regular and irregular shapes out of paper metals or clothing industries. So, all these problems essentially you know; they become very complex and conventional methods of optimizing the finding the best possible solutions often not very you know effective or work very nicely with the conventional methods.

(Refer Slide Time: 02:14)

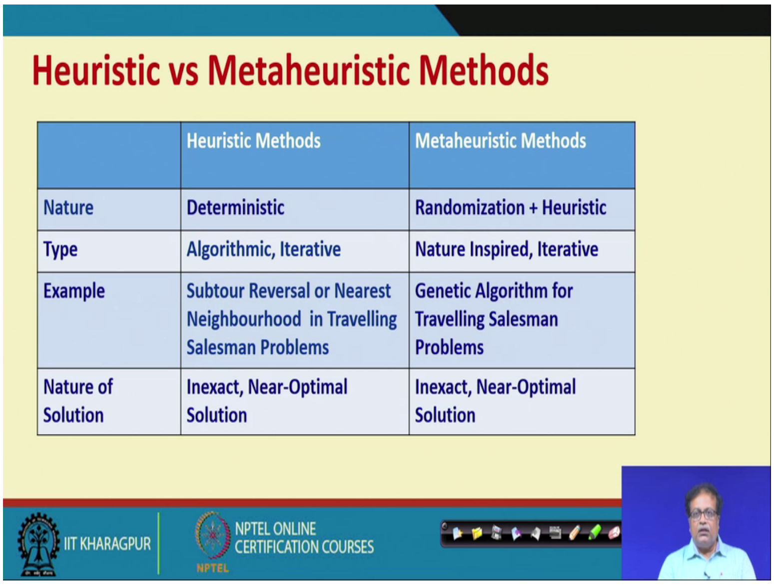


Now, here are the competing techniques what you have say linear or non-linear programming the good point about these techniques like linear and non-linear programming is that they provide exact solutions; the other examples could be your branch and bound which is also gives an exact solutions, but remember the exact solutions, if you really wish to have; you know you have to give quite a bit of computational time to solve large problems, is it alright?

On the other hand, the methods such as heuristic and metaheuristic methods which are essentially some sort of such techniques, you know they are inexact probably can be solved in less time, but they really do not guarantee an exact solution, is it alright? So, they really guarantee a near optimal solution, it can also be optimal solution, but it is difficult to tell whether the solution is really optimal or near optimal.

So, essentially a trade off is required in heuristic and metaheuristic methods between the solution quality and computational time is alright.

(Refer Slide Time: 03:43)

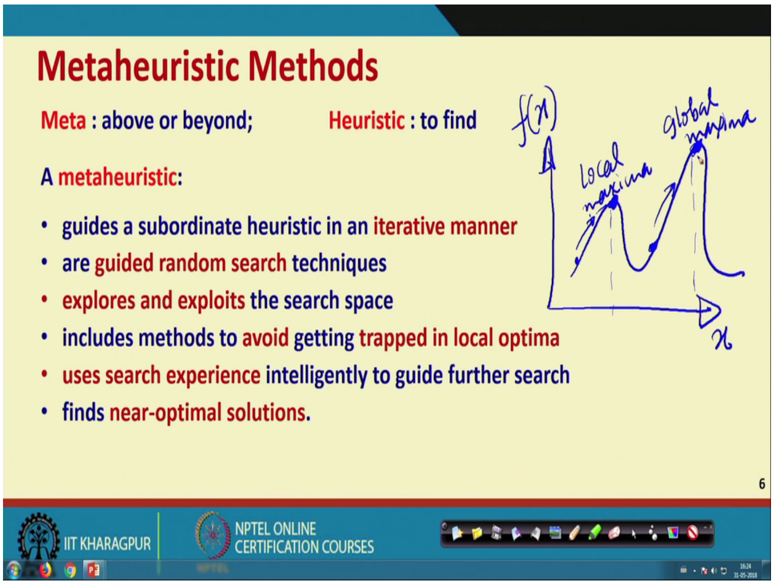


So, let us see; what is the difference between the heuristic and metaheuristic methods; say as you see examples that subtour reversal or nearest neighborhood method that we have employed for ts problems are very good examples of heuristic methods. So, the methods are deterministic algorithmic iterative, but they give inexact or near optimal solutions.

So, what is the difference between heuristics and Metaheuristics? See, the Metaheuristics; they are also use a heuristic along with randomization, is it alright? So, I will I will give examples where it is understood that how randomization is made use of in metaheuristics methods is alright.

Additionally most of these metaheuristics are also nature inspired, right, example genetic algorithms particles on optimization and many such techniques again as you have already said there in exact and have near optimal solutions.

(Refer Slide Time: 04:58)



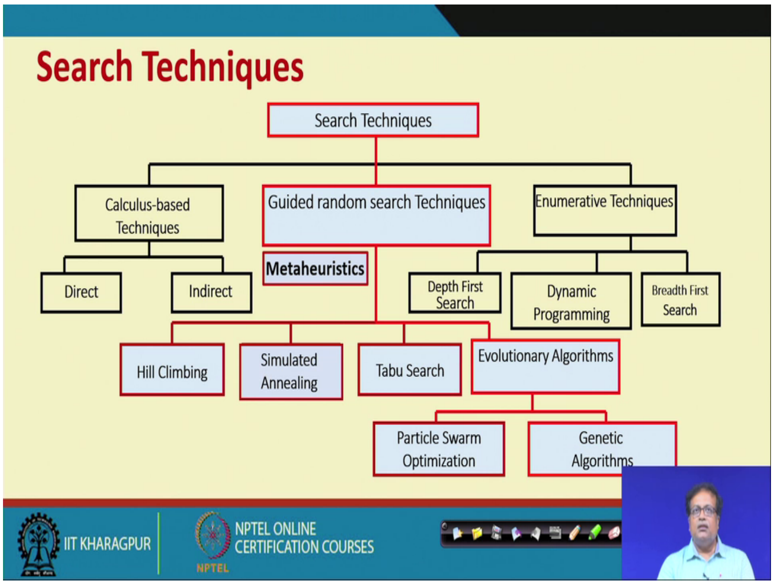
So, look at what really happens in a metaheuristics method, you know, first of all, there is a heuristic which is really subordinate to it and the heuristic the metaheuristics works in an iterative manner along that subordinate heuristic is, alright. So, it is a guided random search technique. So, it is not really starting from single point; it could be starting from multiple points and really does a random search; that means, a random process is there and it tries to explore and exploit the entire search space.

See; what happens, sometimes the function may be having suppose, this side is your x, say suppose, this is the x direction and this side is your functional direction. So, you may have a function which may have multiple peaks and troughs. So, when there are multiple peaks and troughs, say for example, here there is a maximum point here and there is a maximum point here, this is alright. So, this is could be our global maxima whereas, the other point is our local maxima.

So, between the global and local maxima, you know, you can have the function in such a manner. So, supposing the method really starts from a given point here, you know, it goes and it can you know reach the local maxima and gets trapped there. So, this is the difficulty that usually you know heuristic methods are having. So, what really should be done that this trapping should not be done, just assume, suppose, you also start from another point here, then you can reach global maxima. So, between these two solutions, if you compare, then you can know that the solution obtained starting from here gives you a better one.

So, similar such things are possible for metaheuristics methods. So, not only that it avoids the getting trapped into local optima, but it uses the search experience in an intelligent manner and finds near optimal solution. So, that is the essential idea of the heuristic methods.

(Refer Slide Time: 07:53)



So, having known that about heuristic methods, let us look here, you know, this is a kind of you know you can say, the you know the all the search techniques, you know, how do they workout I mean the different types of search techniques, they can be seen here.

So, like you we have what is known as calculus based techniques, then guided random search techniques and then enumerative techniques, right, the calculus based you know techniques could be either direct or indirect. So, what really happens? You know, either you really find out a derivative or second derivative and really obtain through there or use them in a very intelligent manner to have something like steepest descent method or Newton's method and use them like we have studied in our non-linear programming, how to find the best functional value by moving towards an increasing gradient for maxima and decreasing gradient for minima.

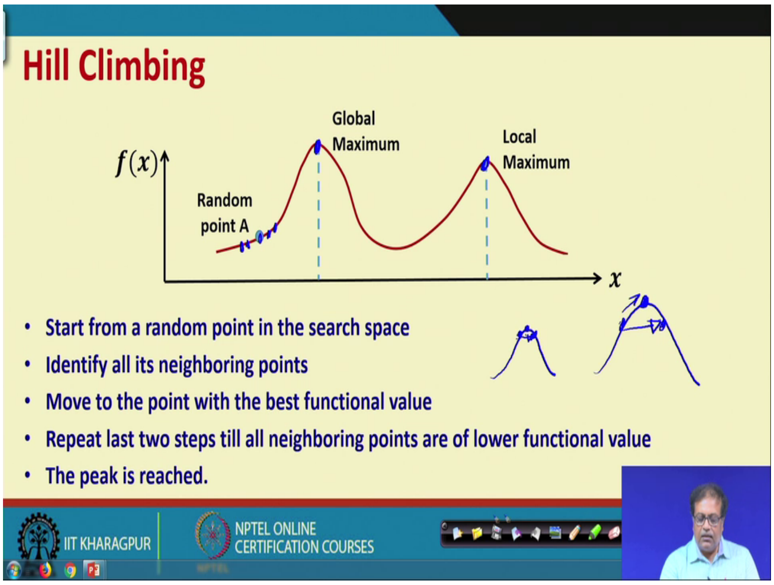
The enumerative techniques again could be depth first or breadth first or you know, we had also studied this dynamic programming methods, you know, they are all can be called another class of search techniques which are really enumerative in nature the on the other hand, we have what are known as Metaheuristics; they are all guided random search techniques and often they are all nature inspired, right.

So, there are many search, but we are going to discuss these once the hill climbing method, the simulated annealing method the Tabu search method, the evolutionary algorithms particle swarm optimization and genetic algorithms as part of evolutionary algorithms, why they are called evolutionary because they happens over generations is it alright. So, we will see the details to understand; what exactly are evolutionary algorithms? The essential difference as you see from calculus based techniques or enumerative techniques, you need some special knowledge about the function such as you know you need the derivatives.

You need the second derivatives or you need some kind of dynamic programming procedures. Now question is all of these really requires the function to be unique, you see it is not possible to really put the find the derivative of functions, all functions, it is difficult to find the derivatives or even when we find derivatives, it is very difficult to really solve the equation by putting the derivative equal to 0.

So, you have seen all that such problems actually are not there in metaheuristics because you know in a very simple language you can tell they are all answered based methods, is it alright? So, if I know that a particular problem has got say one million answers and one or one of them is our best answer, but then may be hundred of them are very good answers or excellent answers near optimal answers. So, how to reach to one of those hundred answers out of the million by a search method; that is what we do in this metaheuristics methods is, alright. So, that is the essential idea of these methods.

(Refer Slide Time: 11:51)



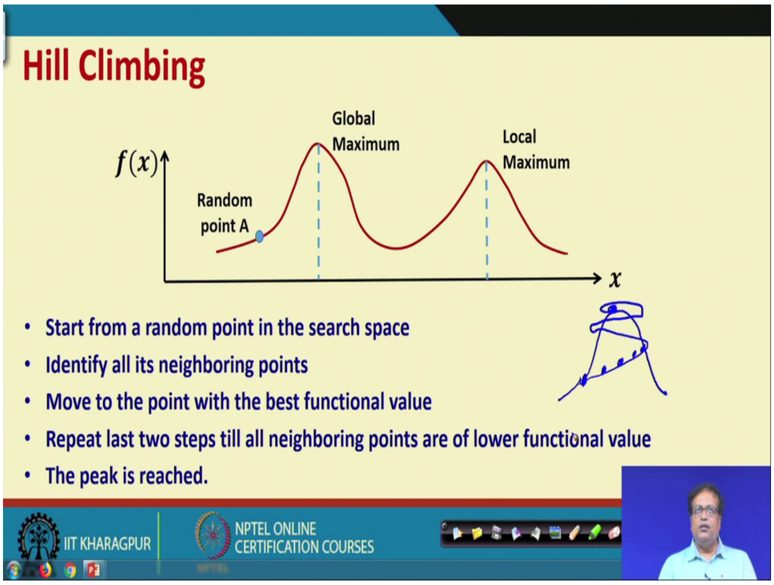
Now, let us look at the hill climbing method. So, what is hill climbing method? You can see here that a particular function, you know has there are multiple peaks and troughs and as you can see this point is our global maximum and this is a local maximum, but supposing we are starting our search from a random point A. Now if we have; now look here, this is only the diagram, I have shown is two dimensional in a three dimensional, this will be a plane is it not and in higher dimensions this curve will actually become a hyper plane.

So, supposing we start from a random point A and then identify neighboring points. So, look here on the functions, there are some neighboring points. So, suppose you choose some neighboring points and then find out what is that point where we are you know getting the best functional value. Now the question is that see there are peaks and troughs. So, you have to be very careful suppose our plot is like this and our point is here. So, this is also a neighboring point. So, if I go from this point to this point not to the peak, then I miss the peak can you see that let me draw slightly bigger. So, this is a point and this is a point and this is a point from these point, if I go to these point instead of going to these point, I have missed the peak, is it alright?

So, you see this point may be higher than this point that is fine, but you know; there are other points which are even better than this point. So, point is when we are seeing the neighboring points we should be careful not to really go too far so that we should not mix miss the really the optimum point. So, we move you know when we say neighboring points it should be within a small region.

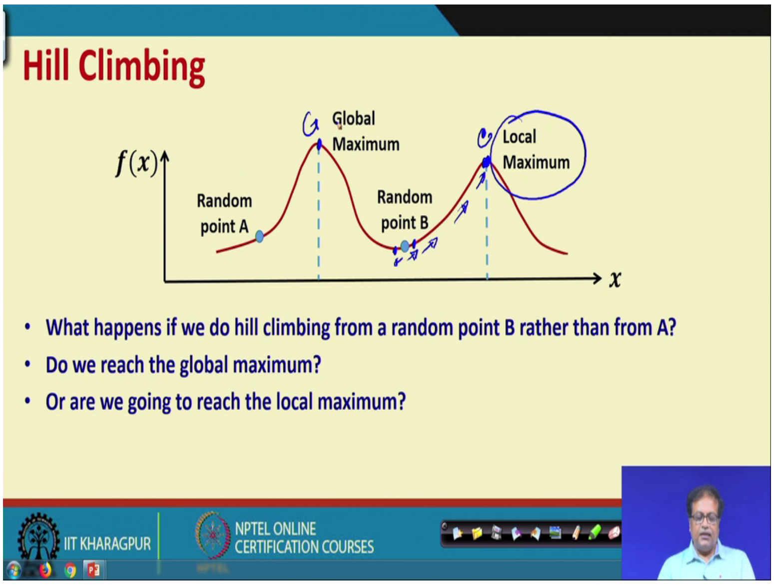
So, move to the point the best functional value and see to it that at any point of time, we really rise through the hill say in a three dimensional hill, supposing, I really have a three dimensional hill like this.

(Refer Slide Time: 14:25)



So, if I have in a point here, I can then move to a point here another point here. So, then like you see in a road, you know, you can you can really move like this really speaking the idea is that at any point of time, if we are doing hill climbing, we should move to a higher point and not go to any lower point. So, that is the essential idea of the hill climbing method. So, keep on moving up and finally, you actually reach the highest point. So, that is the essential idea of hill climbing.

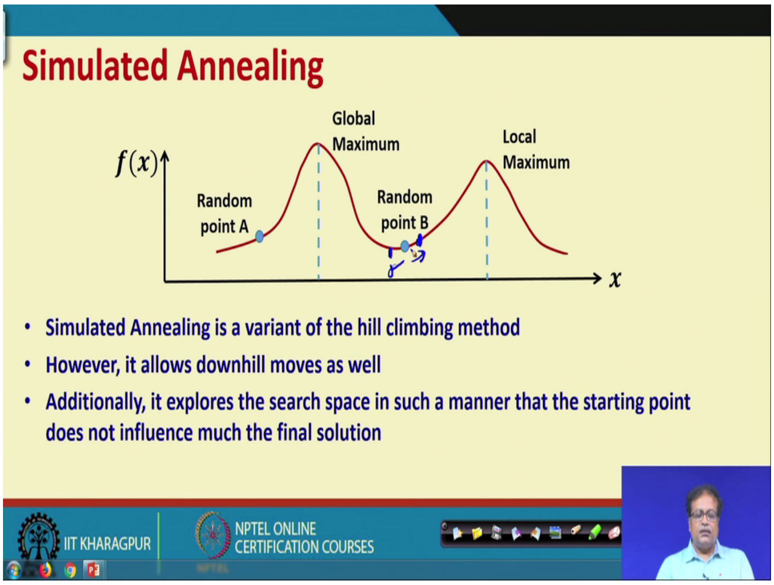
(Refer Slide Time: 15:09)



But then the point should be also remembered that supposing we are doing hill climbing, but suppose we start a point B instead of point A, what will happen there? See we are doing hill climbing. So, if this is my random point B, now any point which are on the other side which are on the other side of point B, say point here and a point here, see, this is going up and this is going down can you see that. So, if I take two points around the random point B, you know, the hill climbing ensures that you move towards a point which is having a better functional value, if we are searching maximum.

So; that means, we actually move up and by moving up you know we finally, reach the point here let us call it say the, but then point C is a maximum definitely, but not the best possible point, it is a local maximum we miss may be point G which is our global maximum; that means, the best possible point. So, this is the difficulty of the hill climbing the hill climbing can actually go to a point which is really not the global maximum, but it could be a local maximum or local minimum instead of the global minimum.

(Refer Slide Time: 16:47)



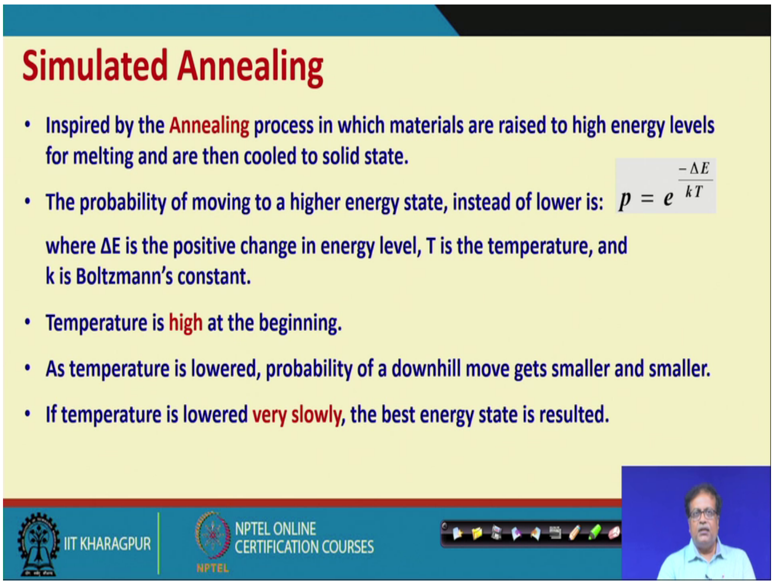
So, these are the difficulties of the hill climbing method. So, what happens in simulated annealing, you know, the simulated annealing techniques you know these specific thing has been looked into.

So, what happens supposing you start from a random point A and then you try to move up hill you reach global maximum, but we have seen for hill climbing, if you start from a point B and do hill climbing you reach to a local maximum. So, what is done here is that you know, it explores the search spaces in such a way that you know a probabilistic value is actually looked into and that probabilistic value is in such a manner. So, that you know, you can also supposing you know like in earlier example, you know, we had discussed about two points which are near the random point B, one is uphill and another is downhill. So, essentially in these method a probability is there to go downhill also, is it alright?

So, through this you know probability, you can actually go downhill also and when you go downhill then; obviously, if you can reach the bottom most point then you can actually move up in the other direction and you can actually reach the global maximum, is it alright? So, this is possible in simulated annealing which was not possible in the hill climbing method.

So, if you are in a point and you are searching for neighboring points this neighboring point search should be such that a probability will be obtained the probability will decide whether to go up or go down, right, both will have a chance and through that chance really the simulated annealing method will also give chance to avoid the local solution maxima local optima solution and go towards a global optimum solution. So, that is the difference between hill climbing and simulated annealing method.

(Refer Slide Time: 19:03)

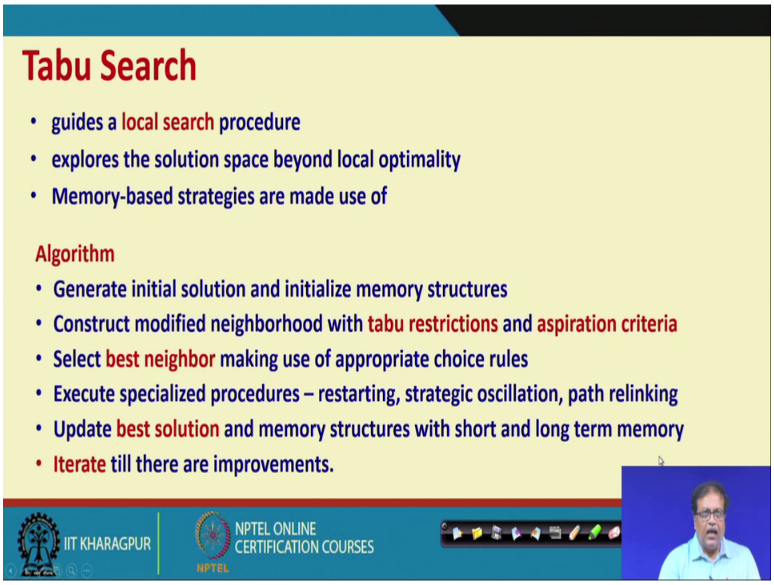


So, as the name suggests that it is inspired by the annealing process the annealing is essentially a process where the materials are raised to high energy level for melting and then it is cooled to a solid state, is it alright, the cooling should be very very low indeed. So, that is the essential idea now the probability of moving to a higher energy state is given by this formula p equal to e to the power minus del E by kT where k is the Boltzmann's constant, is it alright?

Ah, but then that is not the real point here the real point is that as when the temperature is very high; that means, in the initial point then the probability of going downhill you know is rather high; that means, when you have started the iteration iteration process the probability of moving to a higher energy state is something, but there is also a sizable property for going to a lower energy state; that means, going downhill. So, as the temperature is lowered very very slowly the best energy state is resulted, is it alright. So, that is the essential idea of the simulated annealing.

Now, having understood the simulated annealing process we shall discuss then in detail in subsequent lectures, but let us get an overview today about other techniques also.

(Refer Slide Time: 20:47)

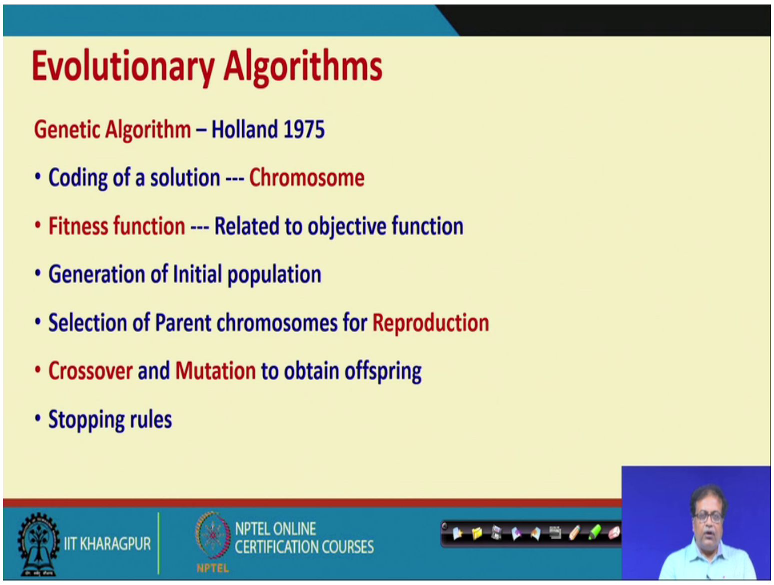


See what really happens in tabu search tabu search is basically a guided local search procedure. So, it explore the solution space beyond local optimality and it uses what is known as the memory based strategies, right.

So, an initial solution is obtained and initial memory structures and the neighborhood is explored by a certain aspiration criteria; that means, which direction to the to be moved, but then also based on the previous results; that means, from the memory supposing we are at a point and we have moved to another point and we saw that that is not a very good direction. So, those directions will be blocked through a tabu restriction is it alright that is that is the essential the tabu means restrictive thing; that means, do not go or do not have a transition from a point A to point B. So, it is a kind of a tabu restrictions.

So, the best neighbor is selected not only by looking at the choice rules, but also looking at the tabu restrictions and the aspiration criteria, is it alright? Some specialized procedures like restarting strategic oscillation etcetera are used for you know choosing the best neighbor and finally, the best possible solution is updated in a continuous manner and iteration is done till improvements are possible. So, that is the essential idea of the tabu search method.

(Refer Slide Time: 22:32)

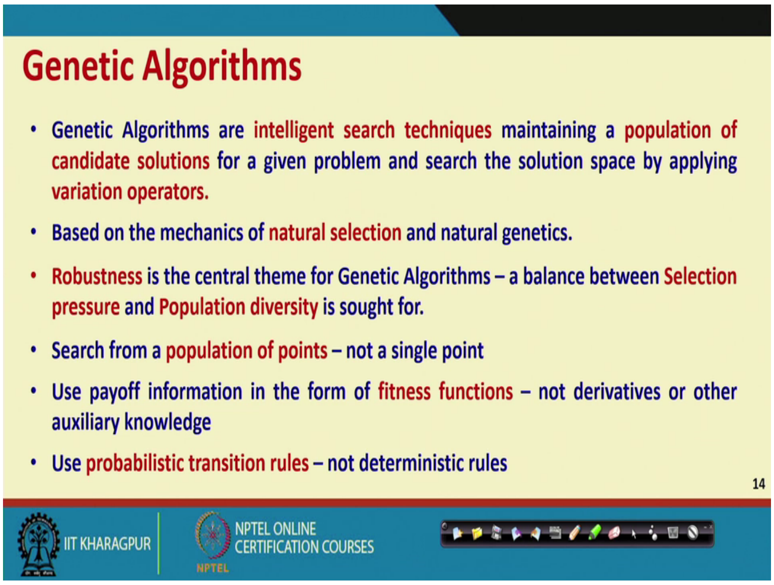


The most important part the major portion of our lectures will really center around the evolutionary algorithms you know, see as I said, these are all nature inspired the nature inspired, here the particular chromosome like in you know the human race or in a any population you know it is a chromosome which is actually holding the what you call the nucleus of every person, alright; that means, the characteristics of a person are all there in the chromosome that the person is having.

So, what really happening here that how good is that chromosome, suppose each chromosome really represents a solution fitness function really evaluates how good the chromosome; is it alright, then an initial population is generated and from this initial population specific methods like cross over and mutation are used to obtain offspring, right. So, you see the first of all the reproduction process starts by the selection of the parent chromosomes. So, which parents will be selected those which are having higher fitness value and then the offspring is generated out of those parents through a process which does not loose the fitness value very much, right.

So, what really happens is that essential idea is from a generation to the next generation the population becomes fitter, is it alright because the parents who are more fit they actually come to the next generation and from more fit next generation I mean more fit parents, we have the offspring who take part in the next generation. So, like this if we move several generations we have highly fit chromosomes or in this case the solutions and; obviously, we have a you know reaching towards the best possible solution by this method.

(Refer Slide Time: 25:05)

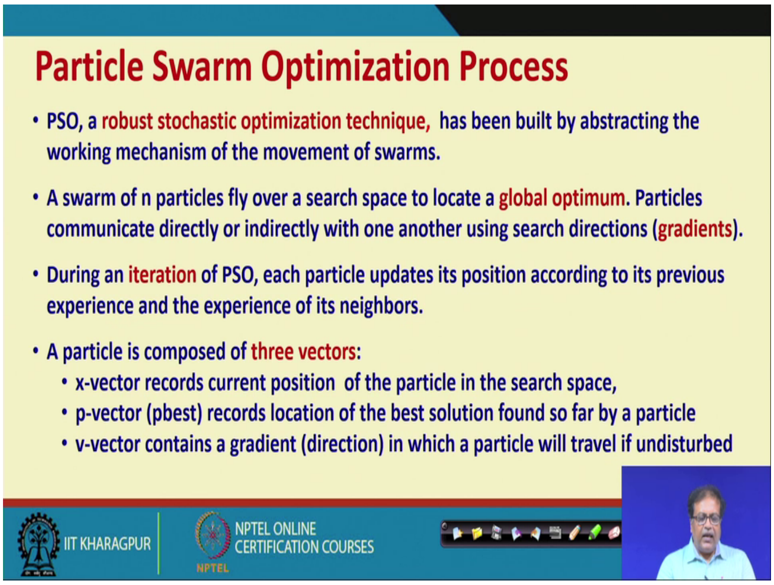


As you know that the genetic algorithms are the most important family member of the evolutionary computing techniques, these are intelligent search techniques maintaining a population of candidate solutions for a given problem which search the solution space by applying various variation operators right. ah

The robustness is the key and the essential idea is that there should be a balance between selection pressure and population diversity you see what is selection pressure the we search from a population of points not a single point. So, if I have a search space, then I start from many points, not just one point, right and which point should be chosen those points which are really having you know higher fitness values, right. So, that is the selection pressure, but at the same time we must really have points from all over the region which is known as the population diversity. So, that we explore the entire population rather than explore only a small portion of the population.

So, advantage that we have is that we do not get stuck in local optima. So, like this, we will discuss more about this it uses probabilistic transitions rules not deterministic ones, right .

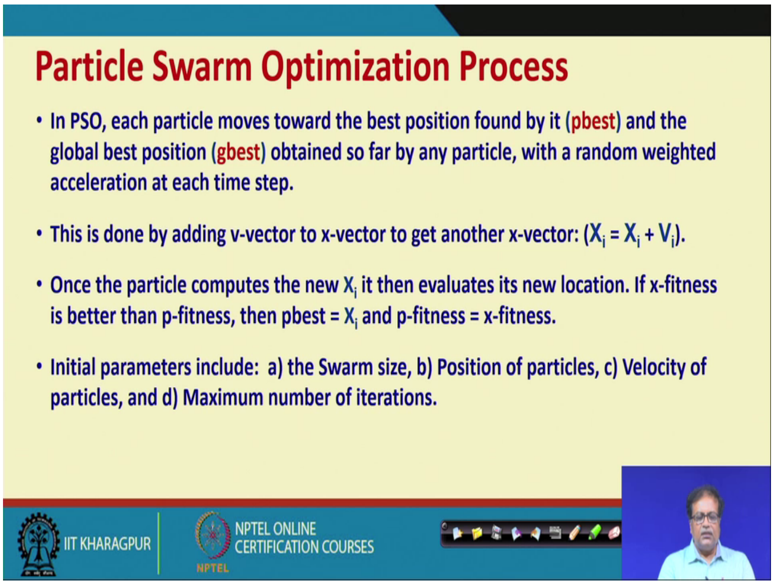
(Refer Slide Time: 26:44)



A particular variant is a particle swarm optimization again this is a robust optimization method swarm of particles fly over a search space to local global optimum and the particles communicate directly or indirectly with one another using search directions.

So, it is like that that we have the entire search space and in this search space we randomly have all these particle swarms. So, different particles are at different points of the search space and then we have vector p vector and v vector and then in every subsequent generations; those particles keep moving rights the moving in such a manner. So, that the you know the fitness or equivalent of fitness really improves, alright.

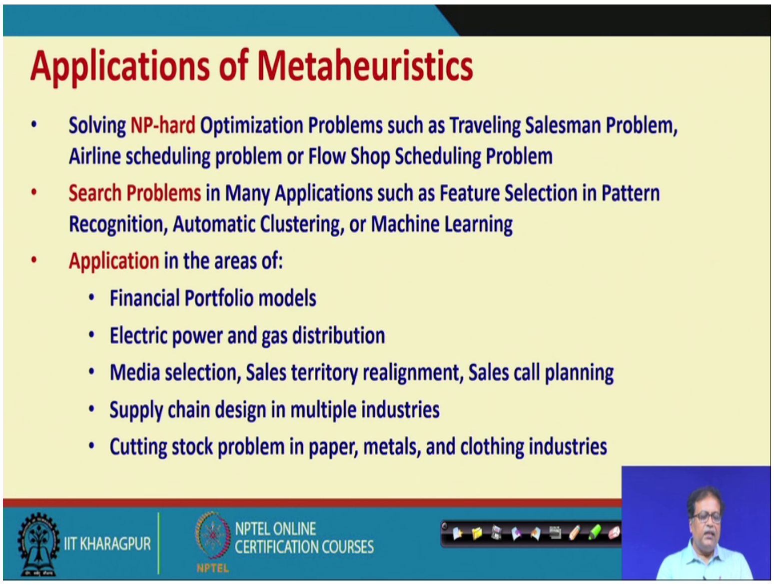
(Refer Slide Time: 27:40)



So, that what happens the particle moves towards the best position that is the pbest and the global best positions that is gbest. So, there are two best one is the gbest another is pbest. Gbest is the global best position of all the particles and pbest is the best position of a given particle, is it alright?

So, a combination of pbest and bbest is made for every particle to make a move in the every subsequent you know generation. So, when all these particles keep moving at a given point what will happen all the particles will really you know move towards the best possible solutions, right. So, that is generally the particle swarm optimization process. So, there are many other evolutionary computing methods like and colony optimization the bat algorithm the bacterial foraging and so on, but we shall really you know keep our discussion concentrated on these many processes that we discussed here.

(Refer Slide Time: 28:53)



Now, there are various applications of metaheuristics methods in solving NP-hard problems like ts problems airline scheduling or flow shop scheduling problems and in search problems in different kinds of pattern recognition automatic clustering or machine learning processes and; obviously, all those application areas which I told right in the beginning, right. So, there are wide applications of metaheuristics and we shall explore this in this particular set of lectures.

Thank you very much.