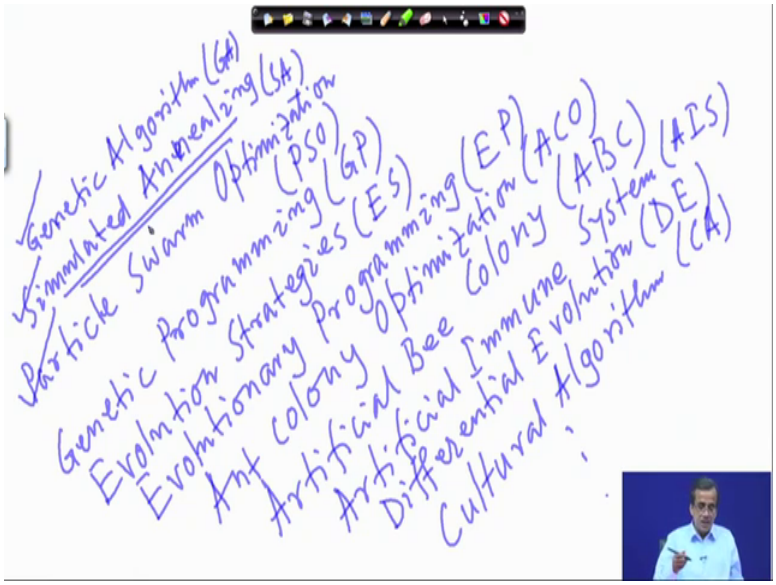
Traditional and Non-Traditional Optimization Tools
Prof. D. K. Pratihar
Department of Mechanical Engineering
Indian Institute of Technology, Kharagpur

Lecture - 20
Simmulated Annealing

Let us start with the discussion on some other nontraditional optimization tool for optimization, now actually we have already discussed the reason behind going for the nontraditional tools for optimization. Now we want some robust optimization tool which can tackle a variety of problems and that is why the concept of nontraditional tools for optimization came and to develop all such tools we take the help of our mother nature.

So, these are all nature inspired techniques, now if you see the list now we have going to large number of tools for example, say if you see we have got the tools like genetic algorithm which are already discussed in details. So, let me just write down we have got Genetic Algorithm.

(Refer Slide Time: 01:18)



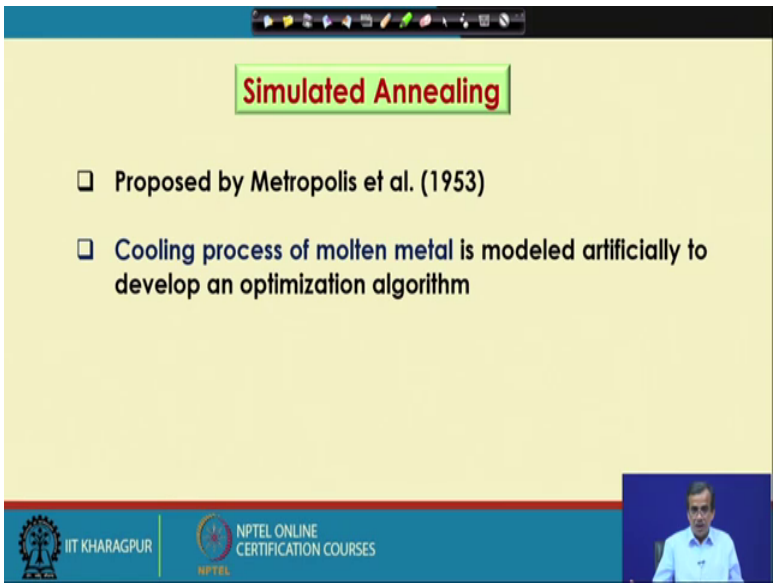
So, in short that is known as G A, then we have got the Simulated Annealing, Simulated Annealing in short this is known as S A, then we have got Particle Swarm Optimization in short this is known as P S O, we have got the tools like Genetic Programming in short this is known as G P.

We have got Evaluation Strategies in short this is known as E S, we have got Evolutionary Programming in short this is known as E P, then we have got Ant Colony Optimization A C O, Artificial Bee Colony that is A B C, then Artificial Immune System it is A I S, we have got Differential Evolution D E. We have got algorithm like Cultural Algorithm in shot C A and many others like bad algorithm and other others, there are so many algorithms now here actually what I am going to do the principal of genetic algorithm I have discussed in detail.

Now I am just going to concentrate on at least 2 more algorithm one is the simulated annealing another is the particles swarm optimization and I am just going to discuss their working principle with some suitable numerical example. Now let me start with the simulated annealing first, G A; I have already discussed, now I am going to concentrate on the simulated annealing. Now simulated annealing that is a topic 9 a on principle actually we are going to solve the minimization problem and actually what I do is, we try to copy everything from a the solidification process of molten metal at the solidification process of molten metal has been copied in the artificial way in Simulated Annealing.

Now, let us see it is principle now Simulated Annealing was proposed by Metropolis et al in the year 1953 long back even before the introduction of genetic algorithm.

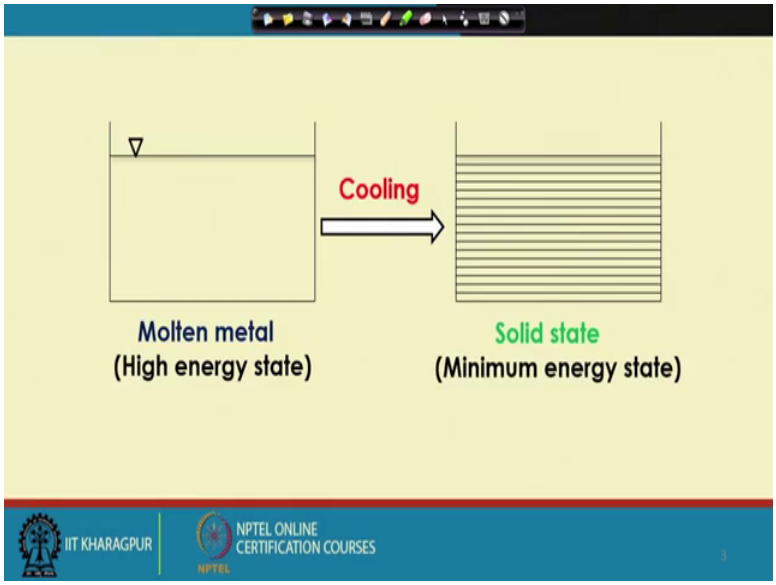
(Refer Slide Time: 05:30)



Genetic algorithm was proposed in the year 1965, but this the concept of simulated annealing came before that in the year 1953. Now as I told that here we try to model the

cooling process of molten metal in the artificial way and we try to develop one optimization algorithm which on principle can solve the minimization problem.

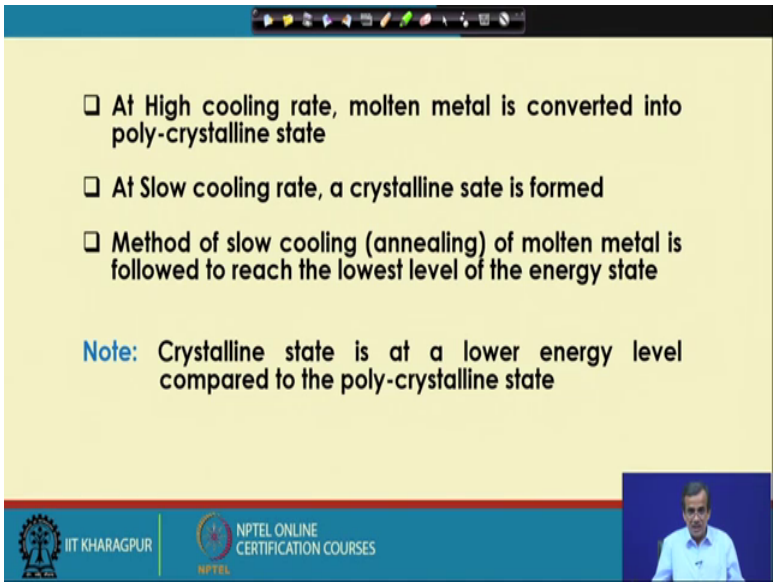
(Refer Slide Time: 06:17)



Now, this schematic view shows so, this is as if the molten metal and this shows actually the solid state of the metal. Now whenever a metal is in molten state it has got high energy. So, this is in a high energy state and if you just cooling down. So, this will become solid and solid state corresponds to the minimum energy. So, from high energy I am just moving towards the low energy or the minimum energy. So, on principle I am solving the energy minimization problem.

Now let us see how to model it in the artificial way and we can develop a tool for minimization, now if you see if you see the molten steel now during this particular cooling if I consider a high speed of cooling or the rapid cooling I will be getting a special type of crystal, but if I use the slow cooling rate I will be getting different types of clusters so, let us try to investigate that fast.

(Refer Slide Time: 07:35)

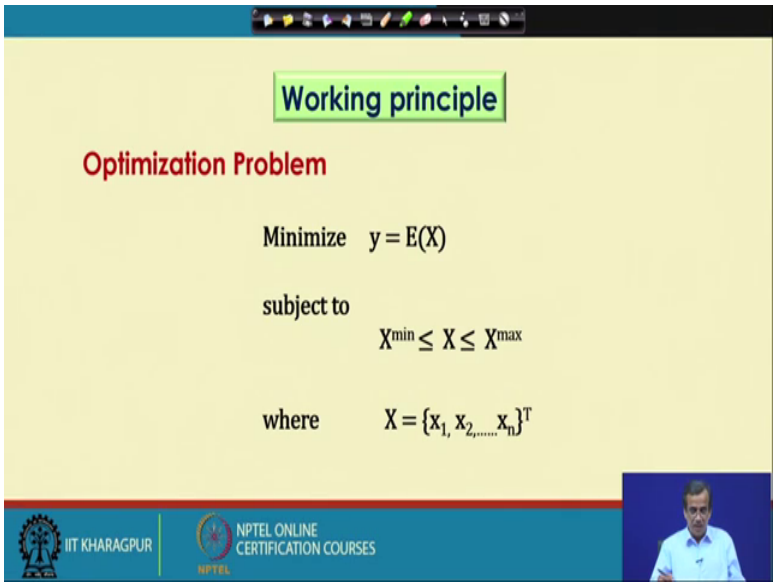


At high cooling rate the molten metal is converted into the poly - crystalline state, on the other hand if I just go for the slow cooling rate so, I will be getting a single crystal a for this particular the metal. But the fact is if I compare the energy level of this crystalline state that is a single crystal state and the polycrystalline state, the energy level of the crystalline state is much lower compare to that of the polycrystalline state.

So, during this particular solidification our aim is to go for the single crystal state. So, that we can get the minimum energy state and that is why we go for the slow cooling and the method of slow cooling is nothing, but the annealing. So, by annealing we mean the method of slow cooling of molten metal and this particular process of slow cooling or annealing has been modeled in the artificial way in the form of an optimization tool. So, I am just going to discuss it is principle in details.

Now, here I have put one note which I have already discussed that crystalline state is at lower energy level compared to the poly crystalline state and we want to minimize the energy; that means, on principle we are going to solve the minimization problem. Now suppose in that I have got one maximization problem what you can do is we can convert it into minimization problem and solve it using simulated annealing.

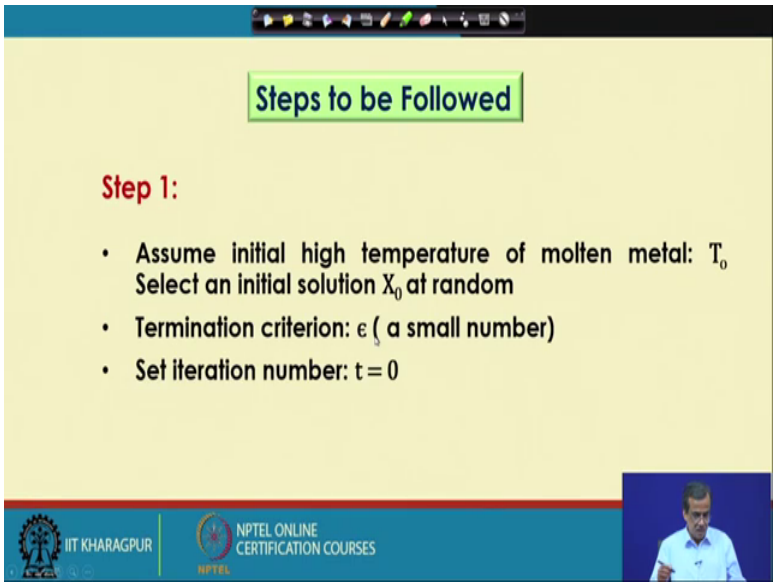
(Refer Slide Time: 09:25)



Now let us see a the working principle supposing that I have got one objective function minimize y is a function of X E X is the energy and the capital X actually it indicates the collection of all the designed variables supposing that I have got n number of design variable.

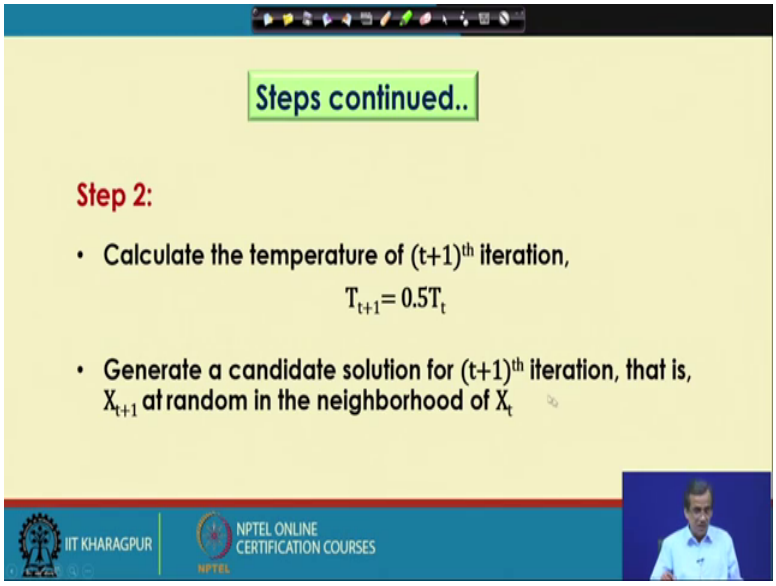
So, capital X is nothing, but a collection of small x 1, x 2 up to x n and this is the range for the variable X is lying between X minimum and X maximum and this is actually a un constrained optimization problem because here there are no functional constraints. Now let us see how to solve this particular the minimization problem using the principle of simulated annealing.

(Refer Slide Time: 10:17)



Now, step 1 we assume initial high temperature of molten metal say T naught and we select some initial random solution say X naught and the termination criteria is epsilon; that means, if the change in energy between the 2 consecutive iteration if it becomes less than equals to epsilon. So, we terminate the program and we say that it has reach that particular the optimal solution and generally we keep a very small value for this particular the epsilon we said iteration number say T equals to 0.

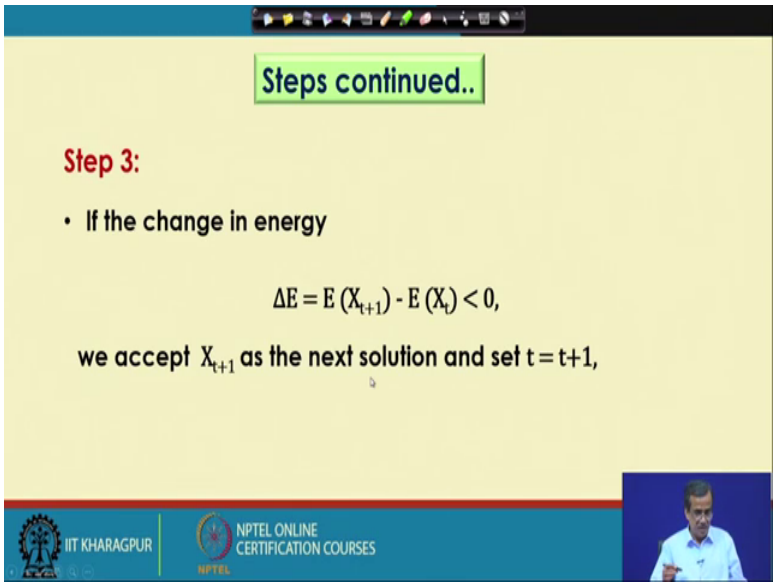
(Refer Slide Time: 11:10)



Step 2 we calculate the temperature of t plus 1 th iteration and that is nothing, but the 50 percent of the previous iteration temperature that is T corresponding t plus 1 th iteration is nothing, but 0.5 multiplied by T corresponding to the t th iteration capital T indicates temperature of molten metal, small t is the iteration number.

So, we generate a candidate solution for t plus 1 th iteration that is capital X t plus 1 as I told capital X carries information of all the small x values that is the design variables and this capital X t plus 1 are selected at random in the neighborhood of X t using the random generator. Now we will have to check whether this particular X t plus 1 should be accepted or not.

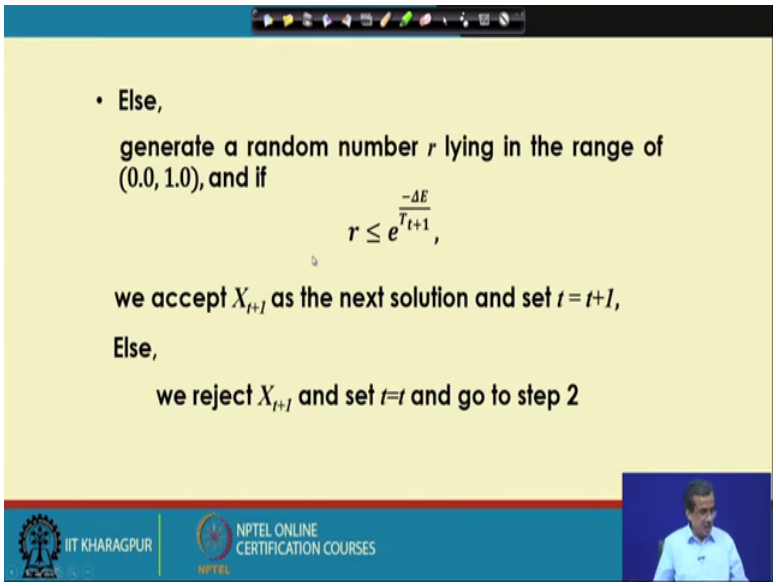
(Refer Slide Time: 12:09)



Now step 3 if the change in energy that is delta E is nothing, but E corresponding to X t plus 1, X t plus 1 I have already selected at random. So, corresponding to that we try to find out the energy and we try to find out energy with respect to X t find out the difference now if this particular difference, now if it is less than 0 we accept X t plus 1 as the next solution and set t equals to t plus 1.

So, we try to find out so, whether this particular E X t plus 1 whether it is found to be less than or greater than compared to E X t plus 1. Now this is a minimization problem. So, if we find that E X t plus 1 is less than E X t that is a better solution and we select that particular X t plus 1 for the next iteration.

(Refer Slide Time: 13:22)



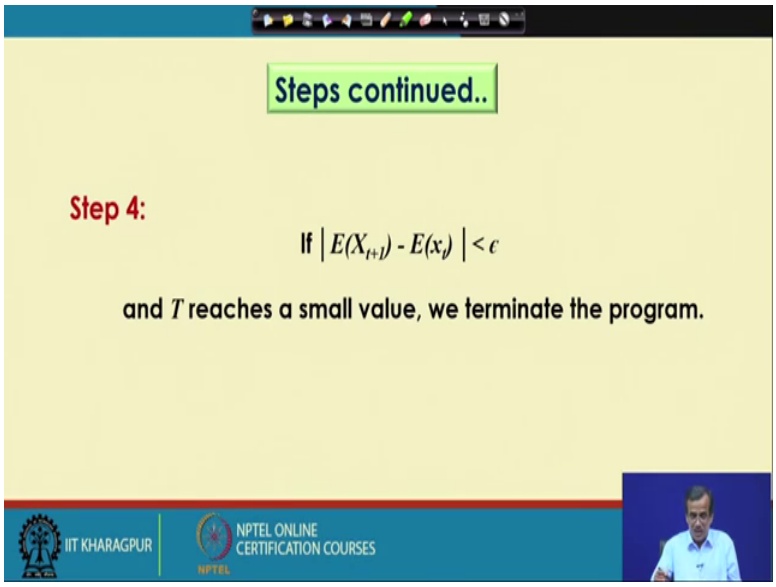
Now, else we generate a random number r lying in the range of 0 comma 1 and we check this particular the condition. Now supposing that the previous condition is not full filled if it is not fulfilled then X t plus 1 should not be selected as the next solution. Now what I do is, if it is not selected through this particular checking now what you will have to do is.

So, you will have to go for the next checking; that means, if this is solution is not selected here we just want to give another chance and whether it can be accepted or not that we are going to check. So, what I do, we generate a random number r lying in the range of 0 to 1 and if r is found to be e raise to the power minus delta E divided by T corresponding to t plus 1 then we accept this particular the solution although it was not accepted previously. So, we just want to give one chance why do we use this type of expression that I am going to discuss after sometime.

Now, if this particular condition holds good we accept X t plus 1 as the next solution and set t equals to t plus 1 else we reject X t plus 1 and set t equals to t; that means, there is no improvement of the solution and we go to step 2, step 2 means we will have to come back here. So, we will have to reduce the temperature generate what should be the possible solution and then we go for the checking once again this if t full fills we select if it fills here we want to give another chance.

If this particular condition holds good we select even if it is rejected previously and if we select this then we go for the next iteration and if it fails here then we reject X t plus 1 and t equals to t and go to step 2.

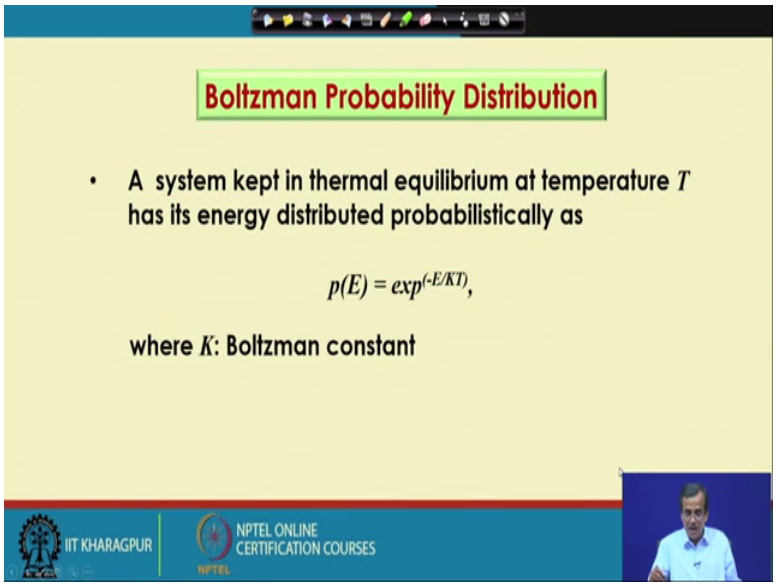
(Refer Slide Time: 15:44)



So, this is actually your the principle, now actually we just want to check the termination condition. Step 4 if the mod value of the difference that is E X t plus 1 minus E x t. So, this will be capital X t yeah there is a small mistake here this should be E capital X t. So, if this particular difference the mod value of that if it is found to be less than epsilon and if t reaches it is small value we terminate the program.

So, this is the way actually this particular algorithm works this algorithm is very fast I have here there is no such operator like G A cross over mutation and that type of thing and compared to G A so, it will be faster. Now I am just going to check and I am just going to discuss, why did I use this type of expression. So, we use this type of expression and there is a valid reason behind that, that I am just going to discuss.

(Refer Slide Time: 16:56)

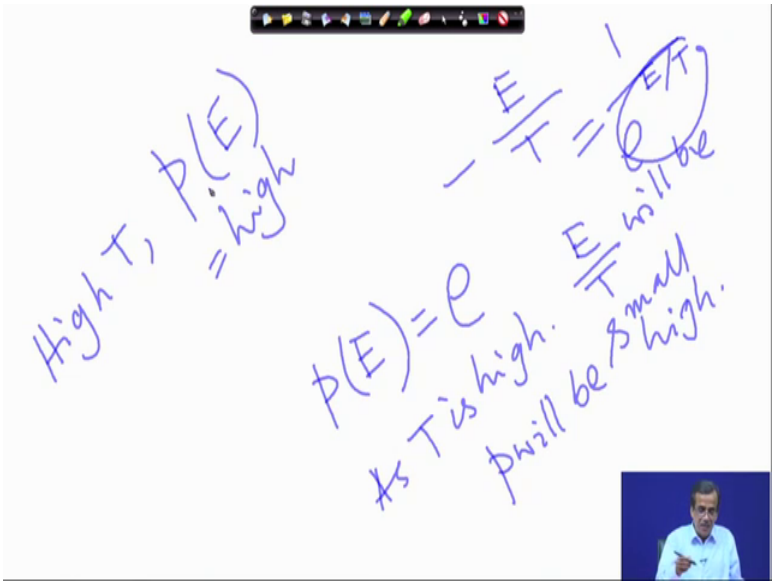


Now to discuss that actually I will have to take the help of the Boltzman probability distribution. Now, according to this Boltzman probability distribution a system kept in thermal equilibrium at temperature T has it is energy distributed probabilistically according to this. So, p E is nothing, but exponential minus E divided by this KT. Now here actually what happens, the system is in thermal equilibrium at temperature capital T and it is energy will be distributed probabilistically according to this particular the rule.

Here the, this K is nothing, but the Boltzman constant and in this simulated annealing what I do is for simplicity we assign a numerical value that is equal to 1 2 K. Now in that case so, our in our simulated annealing algorithm so, probability that energy will be probability of energy distribution is nothing, but a exponential minus E divided by T.

Now let us see, what does it mean, now if I take a higher value for this particular the T. So, higher the value of this particular T so, lower will be E by T because K is 1 here. So, higher the value of T the lower will be E by T and lower the value of this particular E by T. So, 1 divided by so, this exponential that particular E by T will be low and; that means, 1 divided by that particular e raise to the power E by T will have the high value.

(Refer Slide Time: 18:56)

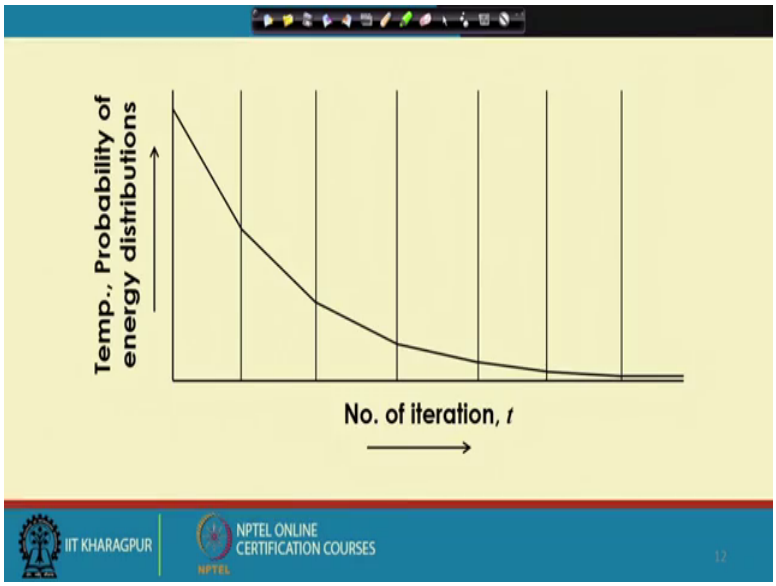


Now let us see how to find out? So, it is nothing, but so, probability of energy is nothing, but e raise to the power minus E divided by T, now here K has been considered as 1. So, as T increases as T is high so, what will happened, your E by E will be small will be small; that means, so what will happen is, if E by T is small. So, this is nothing, but 1 divided by e raise to the power E by T. So, E by T is small means this particular amount is small. So, this probability value will be high; that means, at the high value of T the probability will be high.

So, once again let me repeat at T increases or at high T at high T the probability of selecting the bad solution that is probability of this energy will be actually high. So, if T is high probability is high, on the other hand if T is low the probability is low. Now this probability is actually in simulated annealing the probability of selecting the bad solution.

Now, if I just come back now we can see that so, this is the way actually this particular probability distribution has been considered.

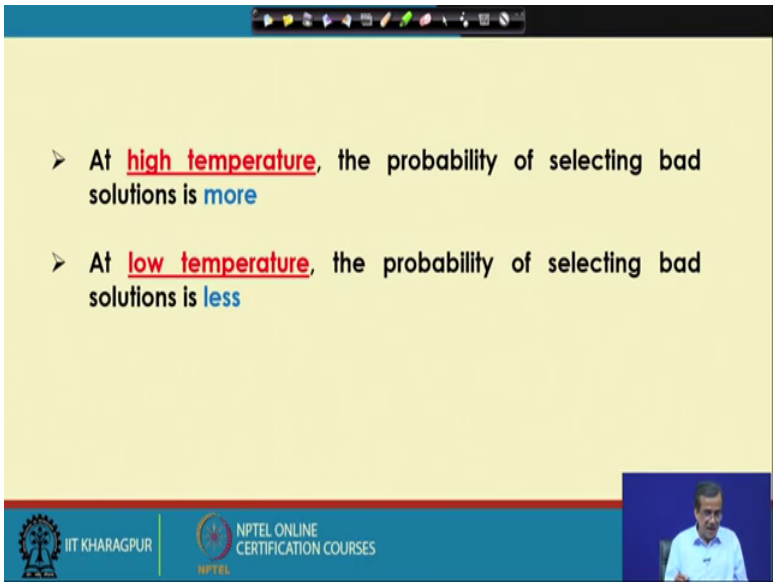
(Refer Slide Time: 20:31)



Now, if you see the distribution as I told when T is high the probability of selecting the bad solution will be more and with the iteration the T becomes low the probability of selecting the bad solution will become small, it means that initially at the starting of the algorithm. There is a probability, high probability that bad solution will be selected , but with the number of iteration the probability of selecting the bad solution should be low; that means, initially we allow mode diversification in the algorithm and with the iteration this particular the chance of diversification will be less.

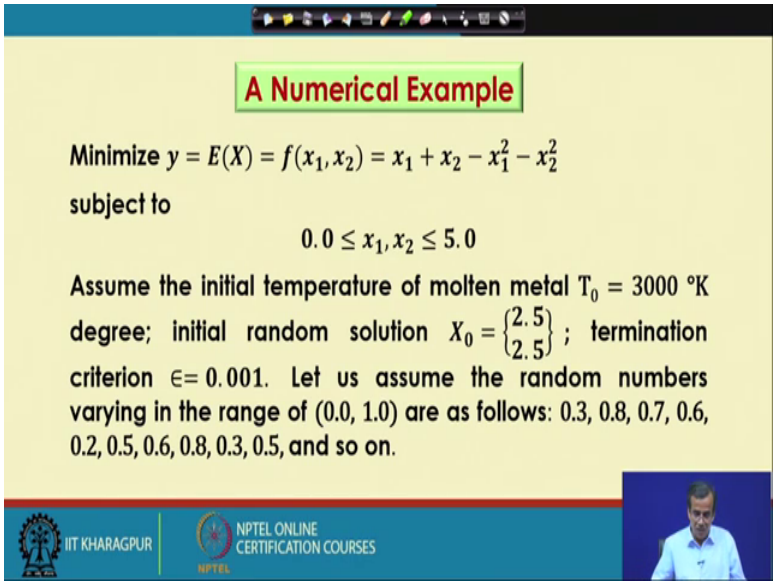
So, this is the way it will try to optimize it will try to minimize that particular the function. So, this is the philosophy behind selecting this type of probability distribution according to Boltzmann probability distribution and this particular thing has been copied in the simulated annealing.

(Refer Slide Time: 21:39)



So, whatever I mentioned same thing I have repeated it here at high temperature the probability of selecting the bad solution is more at low temperature the probability of selecting the bad solution is less and accordingly we have selected that particular your probability distribution.

(Refer Slide Time: 21:54)

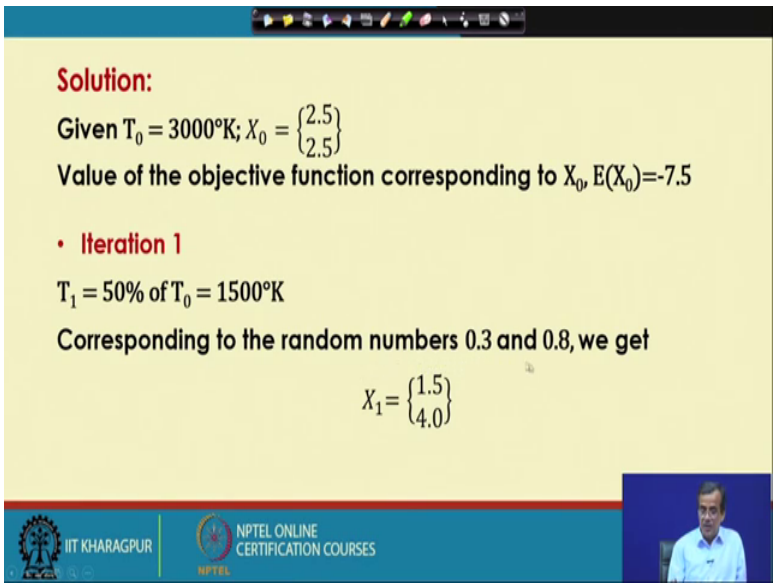


So, this is in short the working principle of a simulate annealing. So, it is very simple much simpler compared to the genetic algorithm and it is faster also. Now to understand the working principle of these particular the algorithm simulated annealing I am just

going solve one numerical example it is very simple. Now the objective function is to minimize y equals to E X that is the energy corresponding to capital X and that is nothing, but the function of 2 variable x 1, and x 2 and that is equals to x 1 plus x 2 minus x 1 square minus x 2 square subject to x 1 x 2 is lying between 0 0.5 and 0.0 and 5.0.

So, this is the range for this particular the design variables and this is once again one unconstrained optimization problem and there is no functional constraint. So, what I do is, we assume the initial temperature of molten metal say T naught say 3000 degree Kelvin and initial random solution we can select lying within this particular range that is X naught is 2.5, 2.5 that is small x 1 equals to 2.5 and small x 2 equals to 2.5 and this is the random initial solution and the termination criteria epsilon is 0.001 and let us assume that the random numbers wearing in the range of 0 to 1 are as follows like 0.3, 0.8 these random numbers I am going to use in this particular the iterations.

(Refer Slide Time: 23:55)



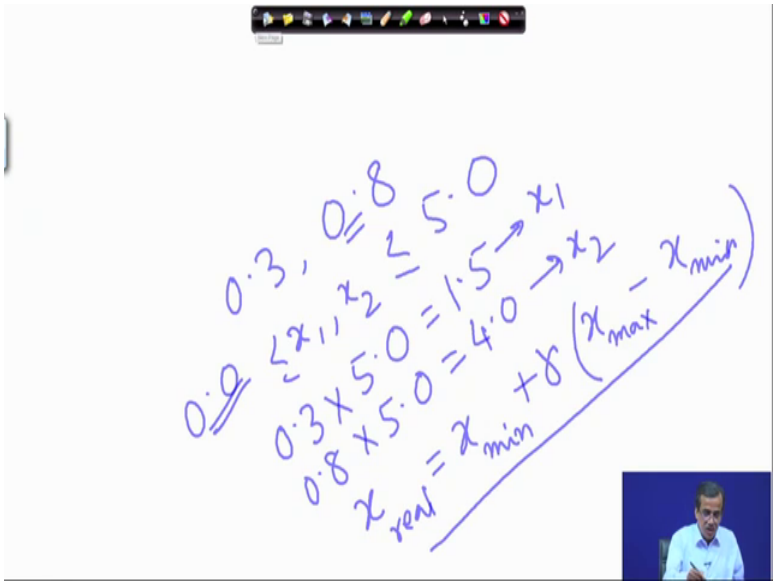
Now, let us see how does it work? Solution: The given T naught that the initial temperature is 3000 degree Kelvin at the initial random solution X naught is 2.5 to 2.5, that is small x 1 is 2.5, small x 2 is 2.5. Now value of the objective function corresponding to X naught that is E X naught is nothing, but minus 7.5, now how to find out this it is very simple.

So, what I do is, you substitute in this particular expression of the objective function you substitute the values of small x 1 x 2. So, will be getting the numerical value and that numerical value is nothing, but minus 7.5. Now what I do is, we will go for iteration 1 so, iteration 1 we find out T 1 that is 50 percent of T naught that is 50 percent of 3000 degree or Kelvin that is nothing, but 1500 degree Kelvin corresponding to the random numbers 0.3 and 0.8 if you remember.

So, we generated a few random numbers here 0.3, 0.8, 0.7 like this. So, I am just going to use 0.3 and 0.8 now responding to these 0.3 and 0.8 the random numbers corresponding to this the real values of capital X 1 that is nothing but the come the collection of small x 1 and x 2 will be as follows.

Now let me let me discuss how to find out this, now the random number which you have considered is 0.3 and 0. 8.

(Refer Slide Time: 25:49)

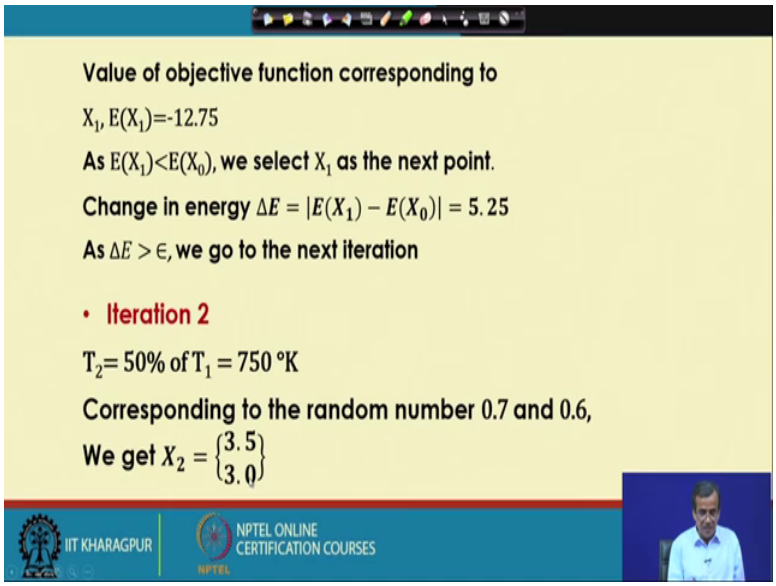


Now what I do is, the range for the variable so, x 1, x 2 if you remember is lying between 0.0 and 5.0, the range is 5. So, what you do is, 0.3 multiplied by the range that is 5.0. So, this is nothing, but 1.5 and corresponding to this 0.8 it is 0.8 multiplied by the range that is nothing, but 4.0. So, I can find out so, this is nothing, but value for small x 1 and this is the value for small x 2.

Actually the rule let me write it here supposing that I know the random number I know the range the rule is as follows if I want to find out the real value for example, x the real value is nothing, but x minimum plus the random number r multiplied by your x maximum, minus x minimum. Now here in this particular problem your x minimum is 0, but this is actually the rule how to find out the real value here x minimum is 0 and random number is a 0.3 x max is 5 x min is 0.

So, I can find out this real value so, this is the way actually we can calculate actually this 1.5 and 4.0 and once I got this 1.5 and 4.0 you substitute in the expression of the objective function and then you will be getting the function value.

(Refer Slide Time: 27:43)

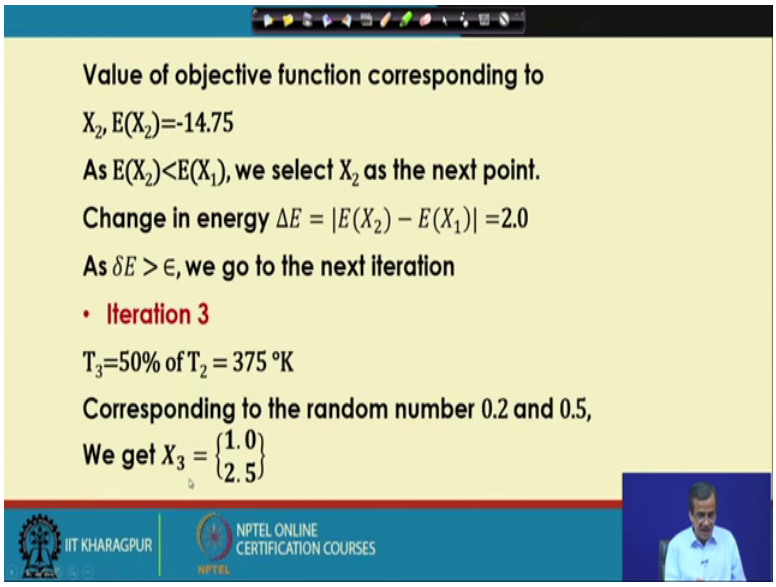


So, you will be getting the function value that is E X 1 is minus 12.75, now we compare E X 1 with E X naught, now E X naught was how much minus 7.5 and E X 1 is minus 12.75 and if I compare we can find out that E X 1 is less than E X naught; that means, E X 1 is a better solution compared to your this thing. So, what I do is you select X 1 as a next point because here E X naught E X 1 is found to be less than E X naught and we are solving the minimization problem. So, X 1 is a better solution compared to X naught.

So, we select X 1 as the next point and then we can find out the change in energy delta E is nothing, but the mod value of E X 1 minus E X naught and here it is coming to be equal to 5.25 and as this particular change in energy is found to be greater than epsilon that is the termination criteria we go to the next iteration. Then iteration 2 so, T 2 is 50

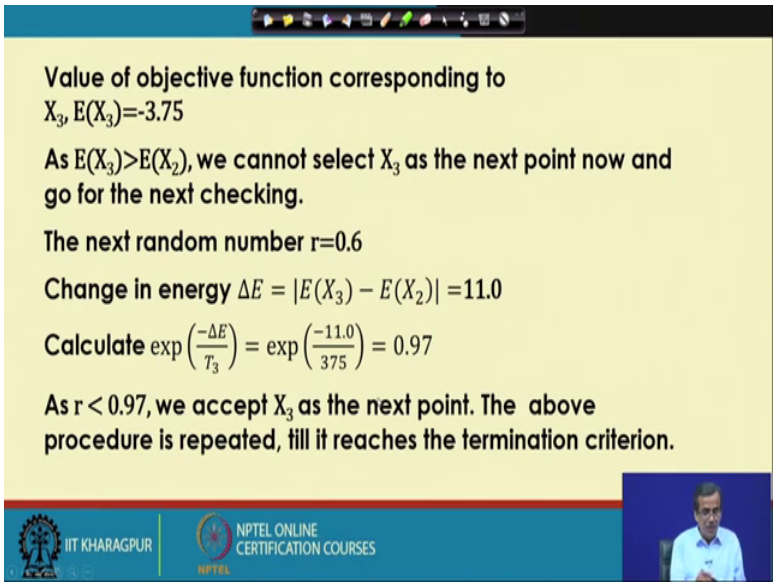
percent of T 1. So, 50 percent of 1.00 is 750 degree Kelvin and corresponding to the random number 0.7 and 0.6. So, you can find out X 2 is 3.5, 3.0 and I can find out the function value is nothing, but minus 14.75.

(Refer Slide Time: 29:21)



Now if I compare E X 2 and E X 1 so, E X 2 is found to be less than E X 1. So, we select X 2 as a next point change in energy delta E is found to be 2.0 and as delta E is found to be greater than epsilon we go to the next iteration this should be delta. Now iteration 3 so, we go for T 3 that is 50 percent of T 2 that is 375 degree Kelvin and corresponding to the random number 0 2 and 0.5 we get X 3 that is 1.0, 2.5 and we find out the function value that is minus 3.75.

(Refer Slide Time: 30:09)



But here if we compare E X 3 and E X 2, E X 3 is found to be greater than E X 2. So, we cannot select X 3 as the next point now and we will have to go for the further checking the next checking. What is that checking? The checking is like this so, which I am already discussed. So, what will have to do is, we will have to find out the exponential minus delta E by T 3. Now how much is delta E is a mod value of the difference between E X 3 and E X 2 that is 11.0 and T 3 is your we know the value of T 3. So, that is nothing, but your 375 so, if we use T 3 as 375 and delta is 11.0.

So, I will be getting 0.97, but the next random number is 0.6 ok. So, this random number are is found to be less than 0.97. So, we accept X 3 as the next point all though it was not accepted previously now according to this particular checking. So, I am just going to accept this particular X 3 as the next point. This procedure is continued till it reaches the termination criteria at that particular termination criteria that is change in energy it is mod value should be less than that epsilon that is the termination value and if it full fills we say that the algorithm has reached that particular level and the solution which we are getting is the optimal solution that is the minimum solution of the objective function.

Thank you