Computational Geometry

Prof. Sandeep Sen

Department of Computer Science and Engineering

Indian Institute Of Technology, Delhi

Module No. # 09

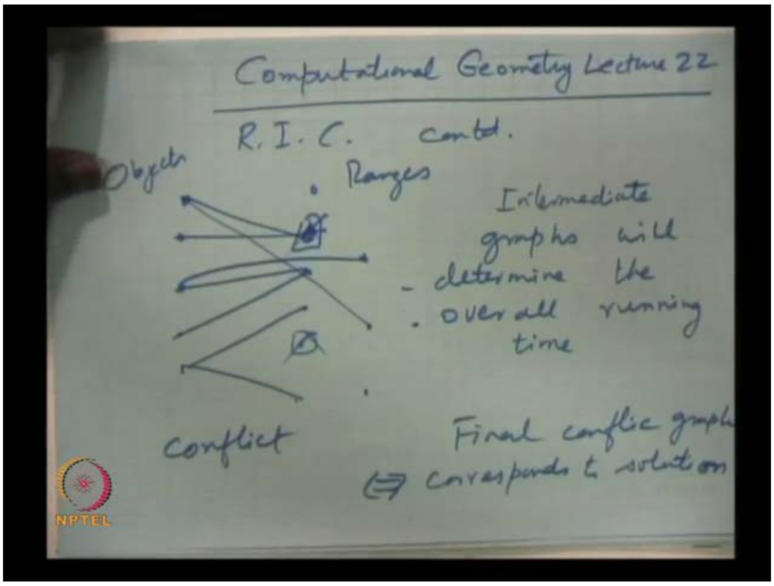
Randomized Incremental Construction and Random Sampling

Lecture No. # 03

RIC continued

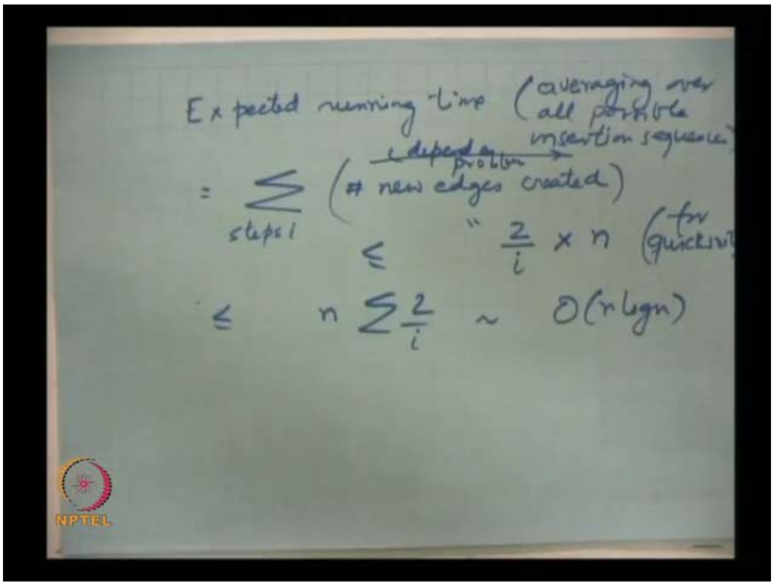
So, we will continue with the with the more general analysis of the randomized incremental construction. So, we started with the example of quick sort. That was in a in a very limited context, where we dealt with only points and intervals. But in general, what I said was, you have this notion of you know objects and ranges right, objects and ranges. So, the randomized incremental construction can be viewed as in maintaining relationship between the objects and ranges.

(Refer Slide Time: 00:45)



And this relationship can be quite general and need not be 1 is to 1 graph. And as you keep on adding objects, you know some of these intervals, you know, they are removed not intervals, I mean in general ranges, and you know some new ranges are created. And there by you are going to add, you know the the new edges are created in the graph, because of these additions, and we keep we keep track of these this changing graph. And eventually, when we have added all the objects, the final conflicts… So, this is called a conflict graph right. So, the final conflict graph - that basically corresponds to the solution of the problem right. So, and all in in and what we obtain as intermediate steps you know. So, that we have to analyze. So, the intermediate the intermediate graphs will determine the overall running time, fine. This is what we observed.

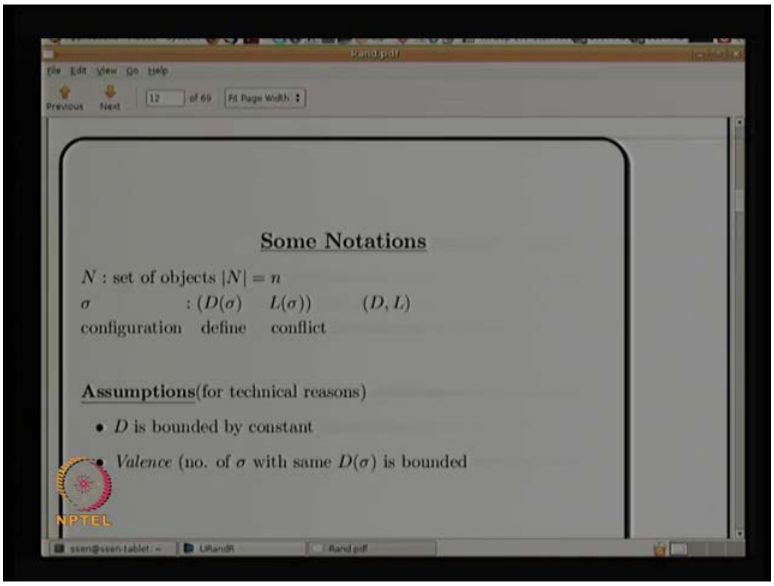
(Refer Slide Time: 02:41)



And the expected running time we argued was proportional to, expected running time which is essentially taken over all possible all possible permutations, averaging over all possible permutation; all possible better seems insertion sequences. That, we argued was equal to the for all steps i, the number of new edges created next line. In the case of in the case of quick sort you know, we argued that this quantity was something like 2 over i for each object right for and since the n objects. So, it this is for we argued rigorously for quick sort right. So, so this. So, this was bounded by this quantity. And therefore, the whole thing was bounded by n times 2 over i roughly about n log n. So, this was how we concluded the analysis of quick sort right. Now, in general I said, the objects in quick sort like points in the ranges was intervals. In general, the objects could be anything the ranges would depend on the context appropriately. So, so let we now, just draw your attention to the notations or some of the definitions that we using to do a more general analysis. So, you want to do a very general analysis of this entire process of you know maintaining this graph, over a sequence of in sessions and at each step will try to when we analyze we will try to bound the number of we will try to bound the number of expected number of new edges created right. that that That is basically be the strategy for analyzing, the general randomizing incremental construction.

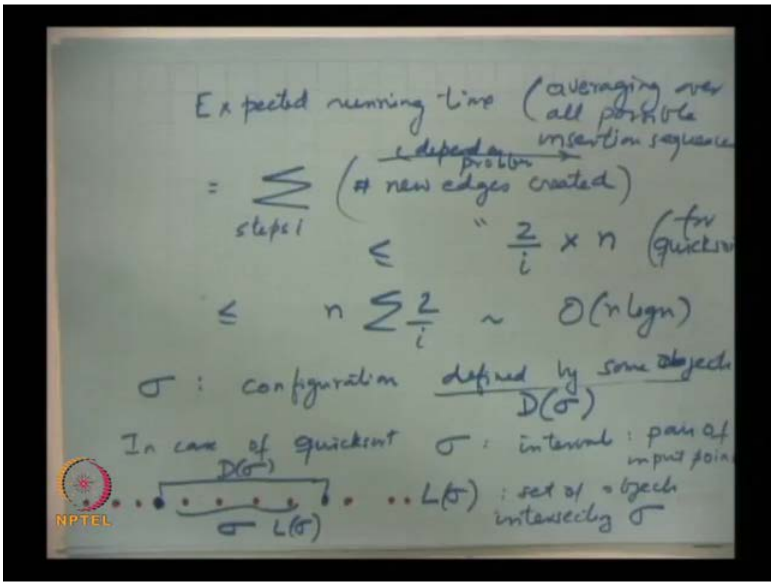
And this clearly should depend on the problem. So, this should depend on the problem - the specific problem. And therefore, if we for for a specific problem, you should be able to just plug this bound, and get the final figure out of it. That is the goal. It is a very general. So, randomized incremental construction itself is a very general construction procedure that can be applied for any you know, all kinds of geometric problems or even non geometric problems, but you know, it has been mostly studied in the context of geometric problems. And then we if you can get a bound, so, this analysis also very general, and if you can get a bound for a specific problem, then we can just you know, plug in that figure and get the final expected running time right. So, what are the notations that we need?

(Refer Slide Time: 05:55)



So, N is a set of objects. So, capital N is a set of objects and small n is a is a cardinality of that. Sigma is what we are calling it configuration.

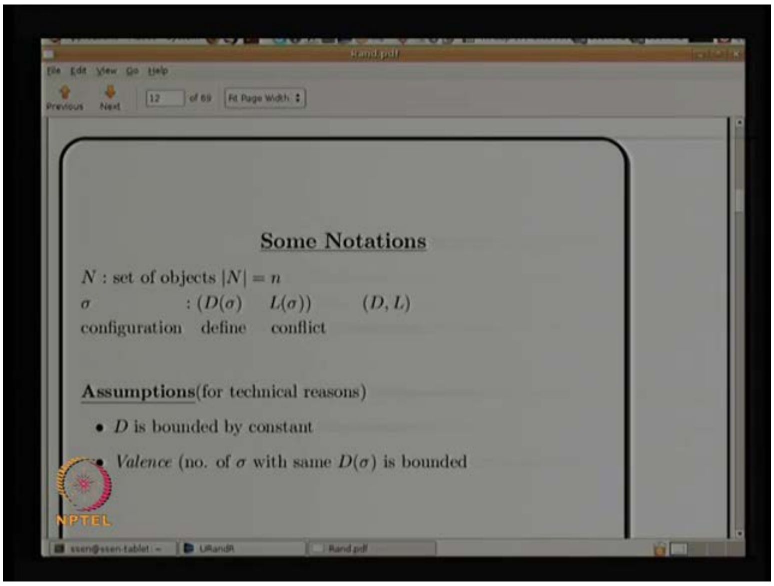
(Refer Slide Time: 06:12)



So, sigma is a configuration right which is going to be defined by some objects. And the set of this objects I am calling D of sigma right. So, in the case of in case of some like quick sort, sigma has a defined intervals - sigma has a intervals, and an interval is defined by the two end points right. So, it is basically pair of points - pair of input points. So, this is defined by some objects. So, suppose this is some sigma, on the other hand, L sigma are all those objects - the set of objects intersecting sigma. So, if you consider all the points on the line, this specific if I am considering this specific interval, then these two are the defining objects of the interval, and the red points basically corresponds to L sigma.

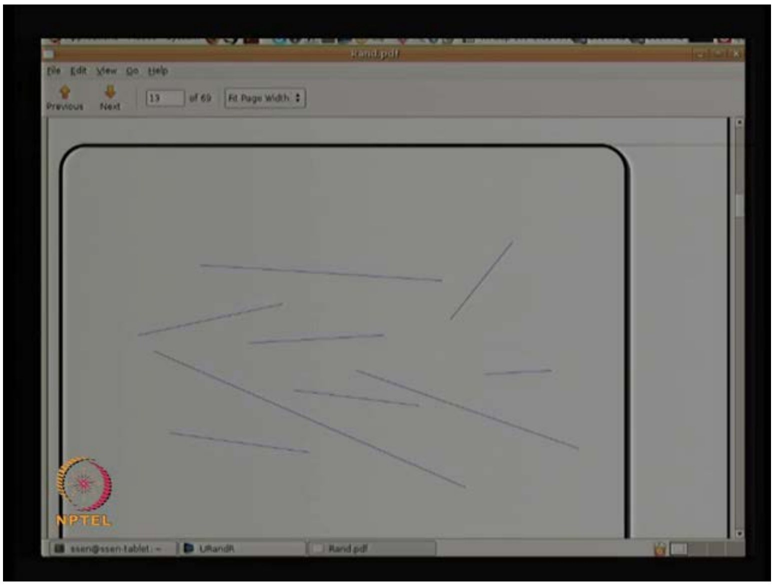
We made this further assumption that D of sigma is bounded by a constant. So, we do not want to get into a situation where an object, the configuration is defined by an unbounded number of objects. So, in the case of intervals it is 2 for we will see some other examples right, may be 3, 4 or 5 you know, but it has to be some kind of constants alright. That is our assumption. And if it is it is valid in most more situations, and the other thing if this kind of technical's. So, let us not worry too much about it.

(Refer Slide Time: 08:23)



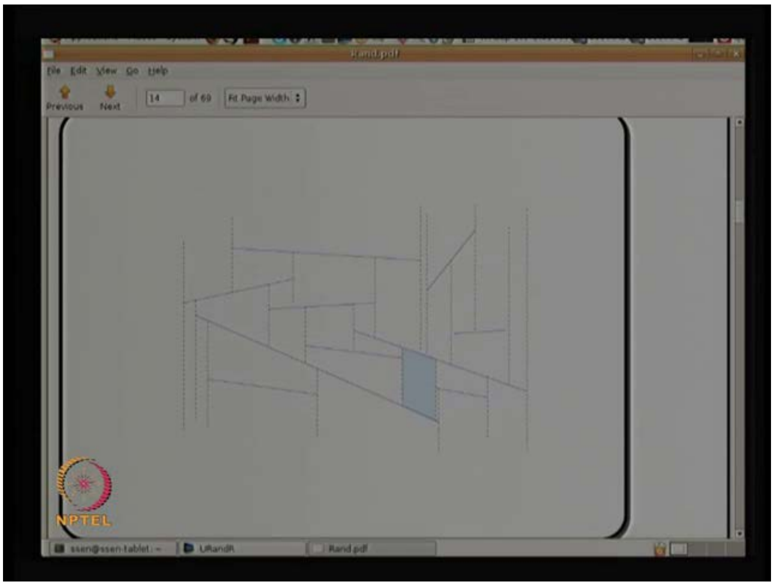
Valence is the number of configurations associated with the same set of objects. So, it is not clear in the context of quick sort, how the same set of points can define more than one interval. But I will just show you an example where this is true, but let us not get into why this is important. So, this is called a valance, and we are seeing that again, even if there is more than one configuration associated with a set of defining objects there cannot be too many such things, again we are seeing the valance that is what we are calling valance and valance itself is a constant, but I am saying this is kind of technical let us not get too much into that.

(Refer slide Time: 09:13)



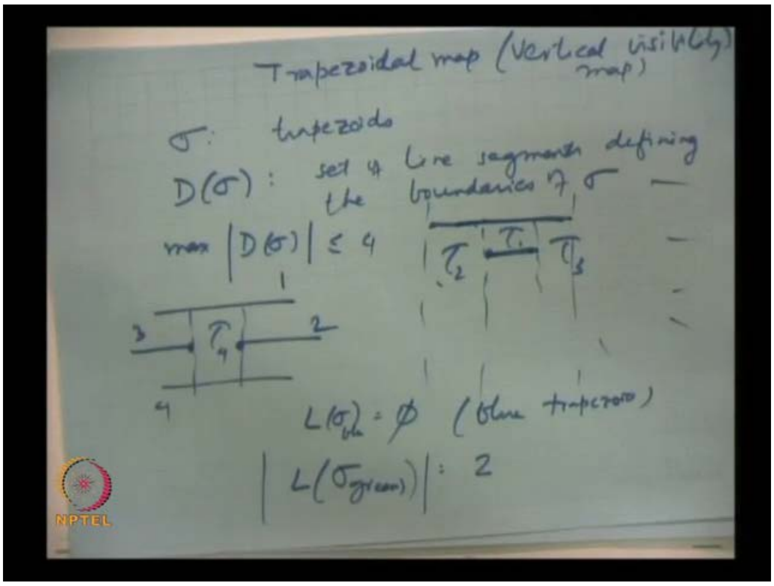
Let me give you another example. This is a purely one dimensional case where it is kind of the kind of prevail we need the definitions. But let us say something like the trapezoid the visibility map problem. The visibility map problem what we have given, a set of non intercepting line segments. And you draw the vertical lines vertical lines to the end points and extend them till they hit something above, but it something below, and those define the trapezoidal.

(Refer Slide Time: 09:29)



So, these are the trapezoidals. So, when I when I extended the vertical lines to the end points. This is what I get. So, this blue shaded figure is a trapezoidal.

(Refer Slide Time: 09:43)



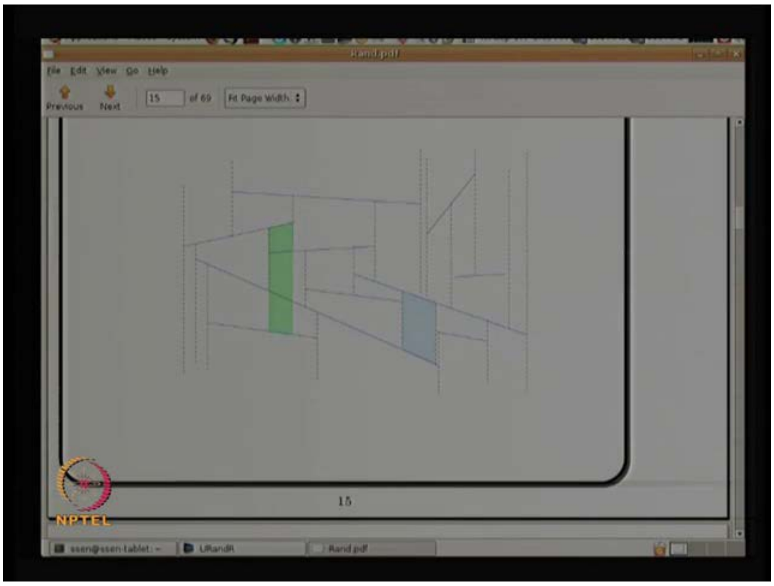
Now, we have already observed that for this problem trapezoidal maps or visibility vertical visibility maps. What we observed is that? So, sigma's, in this case are trapezoids. So, we are for in the context of the problem of trapezoidal maps our sigma's are trapezoids. And so, this sigma is a set of line segments defining the trapezoidal - defining the boundaries of the trapezoidal right.

So, how many or what what what is this thing than the maximum size of D sigma? We actually went through this, does anyone recollect? 4, exactly good. So, we observe that there can be many kinds, you know this trapezoids. So, you know one kind of trapezoidal like this where essentially, this trapezoidal is been defined by this, this and this, it is sufficient. So, it is it is the just these two segments define this trapezoidal, let say these two segments define the the this this trapezoidal, let say let say tou 1. But you can also notice this is another thing, I will draw your attention thus something like this this let us say tou 2 and tou 3, these are also trapezoids, which can be define the same line segment right. May be this are unbounded or something you know, it can be like this. The other segment is a line somewhere here there right. So, you can see. So, this is what I was calling as as a sort of defining as valence that.

There there can be more than one one trapezoid defined in the same set of line segments. So, same set of objects, but there cannot be too many such things. For the same set this these two line segments, we have 1, 2, 3 such possibility. And you can argue that there would not be more than constant number of such trapezoids defined by the same set of same set of line segments. But then again you know, I will I will leave this as a side right now. No, no, let us not worried too much about the valence right.

So, this is one kind of trapezoid and the reason why this is 4. Is essentially, because of you know a configuration like this. So, this configuration is defined by 1, 2, 3, 4 line segments. So, this trapezoidal let us call it now 4 is defined by this.

(Refer Slide Time: 12:28)

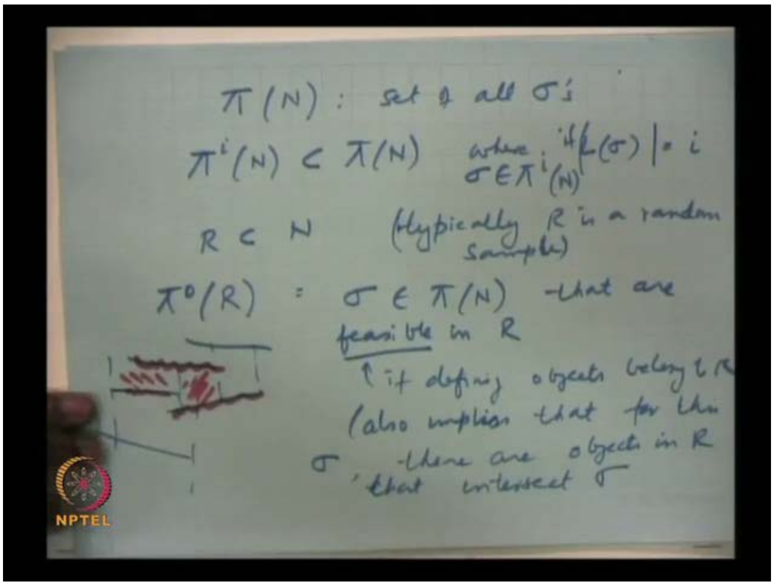


In fact, this blue shaded trapezoid is one such example, it is defined by four line segments. But these are not those. So, the the other kinds of trapezoid, this trapezoid has a particular property. And the property of this blue trapezoid, what what is the difference between this green shaded trapezoid and the blue trapezoid? Here, the blue does not contain any other line segments right. So, so L's of sigma is the set of line segments that intersect sigma, line segments will object the intersects sigma.

So, in this case the blue, if if you look at the blue shaded trapezoid, L sigma is basically empty, you know, for the blue one. And whereas, if you look at the green one, so, let us call this L sigma blue and L sigma green. So, that has it has two intersecting objects or intersecting line segments. So, L's of sigma, the cardinality of that is - that is 2. I will skip. So, this is an example for Voronoi diagram, but for the time being I will skip this example. Why is it 2 and not 3? That is good, why do you think this should be 3?

(Audio not audible. Refer Time: 13:49) the number of objects that are intersect it. Not, not leaving aside the defining object. Not including the defining objects. So, I will skip this example of Voronoi diagram.

(Refer Slide Time: 14:17)



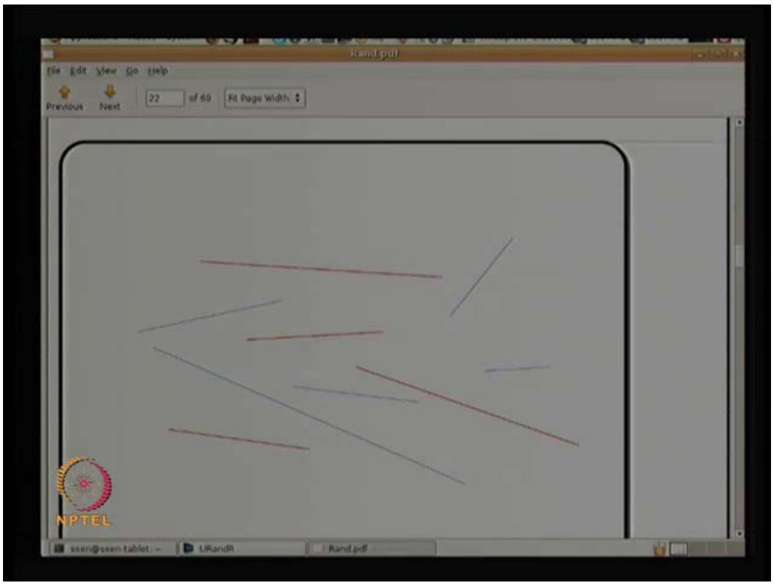
So, continuing with the definitions. So, what is being said here is this pi of N is a set of all configurations for given set of objects. So, size of N is the set of all sigma's. So, these sigma's may be like the green one or the blue one, it may be empty, it may be non empty whatever. So, this entire set we are defining as pi of N. On the other hand, pi i of N is a subset of pi n, where L the… So, sigma sorry, where sigma belongs to pi i of N, if else the size of this conflict - conflicting objects is equal to i. (No audio. Refer Time: 15:15 to 15:29) So, in in the previous example the green trapezoidal, the green sigma would belong to pi 2 of N right, whereas the the the blue one was you know pi 0 of N, because you know the no conflict this is.

Now, R is a sub set of N R is a sub set of N which is let us say in most cases, you know, typically R is a random sample R is a random sample. So, then pi 0 of R all those sigma's that are feasible in R. So, what is feasible in R? If the defining objects belong to R, only then it is feasible enough. If I take a subset if I take a subset, so, I have these you know, these lines line segments there many such trapezoidals defined by this. If a pick a random sample, suppose, sample any any subset basically. So, then this particular trapezoid is feasible enough, because the defining segments also belong to R. But something like this, this is not feasible enough, because the defining this for this trapezoidal this was one of the defining segments, this does not belong R. So, this is not feasible enough.

So, some of these trapezoids will be feasible in the subset, only if the defining objects are there. And more over, since I am writing 0, it will imply also implies that is the same thing, like it was pi 0 of N, this is pi 0 of R. So, inside R… So, implies that for this sigma, the there are no there are no objects in R that intersect sigma.

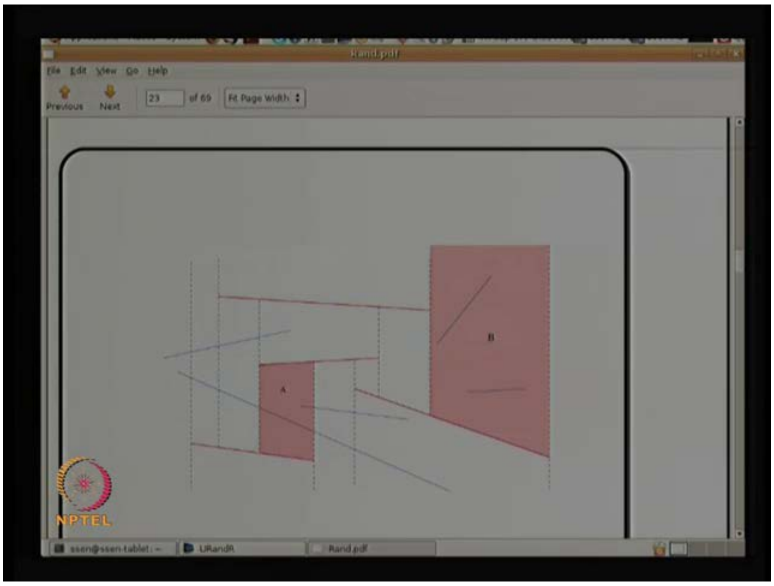
So, what are we defined, you have basically defined a configuration which is defined in terms of some of constant number of objects. For a a given configuration sigma, what are the defining objects, what are the conflicting objects, D's of sigma, L's of sigma, pi's of N is a is entire set configurations, and then pi of N are those configurations for which L sigma has size i. And whenever we restricted to a subset, then we have a notion that you know certain the certain configuration is active only the defining object are in that sample. These are the things that these are the only these are the only things that you need to basically remember, but you know in the beginning, you know, you may have some you know, there will may be some confusion with these notations.

(Refer Slide Time: 19:06)



So, here is an example. So, the blue, the red segments are basically R. And see, if you look at A, if you look at the trapezoid A, this trapezoid belongs to pi 0 of R, red is the is the sample set. It has no conflictive objects that belongs to R, but it has conflicting object that belong to the original set. Why are we doing it?

(Refer Slide Time: 19:15)



One reason that you know, we are trying to define all these things is, if one watt in some very generic sense of very general sense you know, construct, let us say the trapezoidal map is in some kind of divide and conquer. So, in two dimensions it is not clear, you know how do actually divide the problem. So, one possibility is that we pick a sample, and based on the sample we define these trapezoids, and when and these trapezoids will basically be the sub problem why the trapezoids that belongs. Why? The trapezoids that belong to pi 0 of R, they will have some conflicting objects from the original problem, and then we can actually solve the problem restricted to the trapezoids. So, like we can solve recursively all the segments that intersects trapezoidal. And it is it is a visibility vertical visibility problem. So, we are not losing any information. So, within that trapezoid, you know we will do some kind of recursive recursive call, and then what what would you really what desirable what is desirable is that for divide and conquer that the sub problems as I should be too much. So, if I pick a sample of let us say size R, ideally I would like to have a sub problem of size N over R.

So, that is easy to visualize in the case of one dimensional problem. I pick up hard samples, if that pick them if I knew if I pick them exactly equally apart there will be N over R points in each of the intervals. It is not at tall obvious in the case of a high dimensional problem that know, what is those exactly for you know, what there is no ordering. So, the reason why it is so simple in in one-dimensional, so natural, because there is a natural ordering in one dimension. For a two-dimensional problem that trapezoidal maps, you know, we do not even have a notion of ordering, how do you order two things in two dimensions as a high dimensions. But we still we are still trying to define some notion of dividing the original problem into some sub problems, such that each sub problem should have a bounded size. Bounded means you know something if I pick is, you know, if I want to divide into R parts roughly each sub problem should have N over R sides.

In one-dimensional, if I corresponds naturally to the link (( )) in this case (( )) corresponding (( )) in this case.

I mean in the size corresponds to what length; we are not talking about any lengths, we are talking about the number of points in an interval.

(( )) number of points.

So, here exactly. So, now, L sigma is basically the sub problem. If sigma is a sub problem, if sigma belongs to pi 0 of N, pi 0 of R is a sub problem, then the natural thing is that all the all those objects that intersect sigma defines the sub problem. So, this two blue this two blue segments that intersects trapezoidally is precisely a sub problem right.

So, now there is a difference between A and B, can you sort of tell me what is the difference both of them have some conflicting segments, but there is a difference.

(Audio not audible. Refer Time: 23:04)

Exactly. So, the sub problem B has segments, you know that only live in B, whereas the sub problem a, if you are defining that trapezoidal sub problem that has this two line segments which also intersects other trapezoids. So, this is the case that I was I was pointing out that in in the case of when when we are analyzing this conflict graph, each object could belong to only one interval, in the case of quick sort. But here, if the it is the objects are now line segments. And if the sub problems are trapezoids or this sigma's, then a single object can conflict with more than 1 sigma's. So, it can belong to more than one sub problem. So, you are not in that very simple situation makes A 1 to 1 mapping, each object belongs to only one sub problem. Now, in trapezoid B of course, there is a nice case that these two belong to only B and nothing else. But in A there are these objects that may be belong to more than one trapezoid. And so, what happens, as a as a consequence of this what happens, when you do a a recursive call what happens.

(Audio not audible. Refer Time: 24:20)

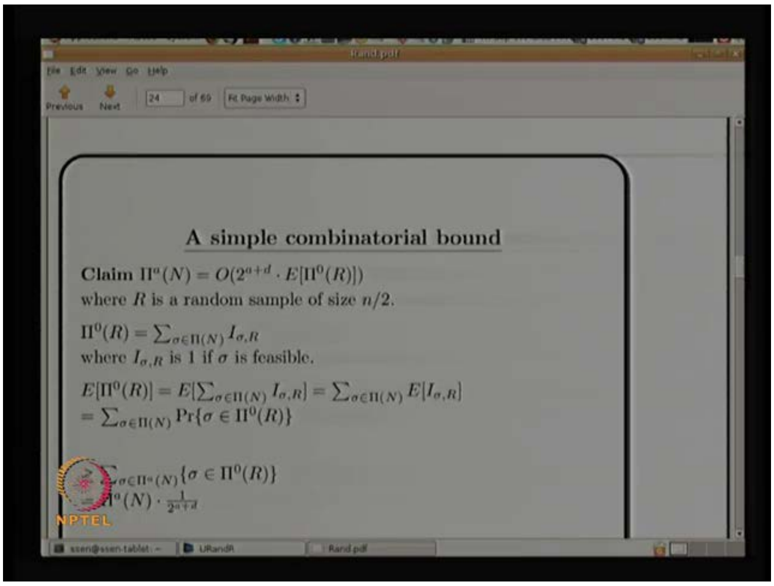
No, it is not overlapping problem, see I can I can always clip this line segments to within the trapezoid that is not a problem. (Audio not audible. Refer Time: 24:29 to 24:34)

Exactly. So, what is happening is that? If you look at, so, one one aspect of divide and conquer that we normally tackle is a size of a sub problem, but the other aspect is what is the total size of the sub problems. So, kind of assume that equal to n, but you know in in a problem like this or the way that we are trying to define the sub problems, you can no longer assume that the total some of the sub problems is equal to N right. So, it can increase, because each object can conflict with more than one trapezoid, and when you sum them up, you know it is going to basically, you know it it can easily blow up.

(Audio not clear. Refer Time: 25:29)

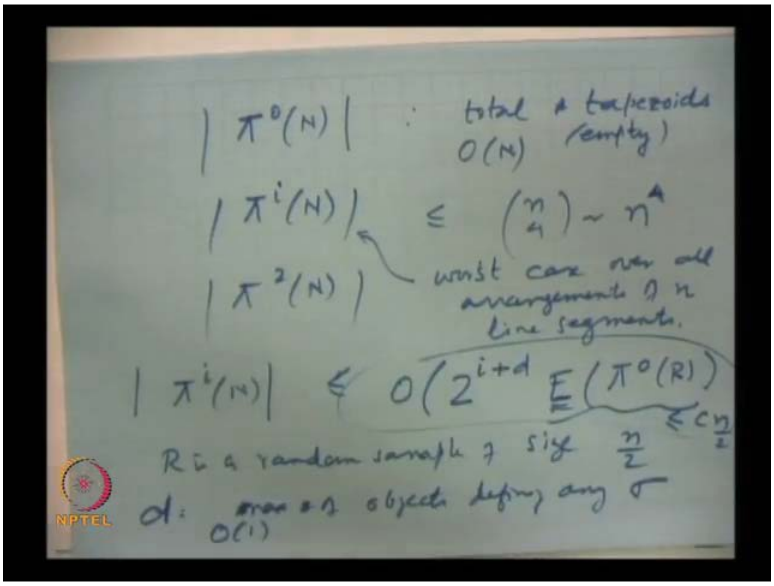
Right, right. So, that is what I am point out, but you know, you do not encounter these things, you know when you are looking at this points and intervals and you know one-dimensional problems. But here, you know, there is a clear cut case for this, you know and for many other problems, we will we will have this situation that the same object can conflict with many other; many many configurations and can belong to more than one sub problem. So, there is a potential blow up as you go down in the recursion, which can cause, which can cause in efficiency right.

(Refer Slide Time: 26:08)



So, let me just give you, very nice illustration of you know how random sample can be used to prove combinatorial count. So, I am not even talking about algorithm, you know.

(Refer Slide Time: 26:35)



So, suppose I was interested to count how many… So, so this, this happens to be very important combinatorial problem. So, in the case of trapezoids, you know, you can look at this quantity pi 0 of N, what is the size of this that is a total number of trapezoids right. There is a total number of trapezoids. So, these empty trapezoids right; suppose I was interested in this quantity; can you count this the first one? So, how many trapezoids can you create empty trapezoids, when you draw do this vertical visibility map. Say it again, some odorant clearly it is a planer map you know when we argued, in fact, I gave with figure like six times, you know three times something like that.

So, this is clearly odorant. So, we know this basically from the fact that it is a planer map can you count this quantity? I am interested to know the number of trapezoids; you will get a bound of the number of trapezoids that have I intercepting segments; what is the total number of trapezoids without bothering about i?

Right so four segments defining a trapezoid, the maximum number of trapezoid is the N to the power 4. So, a trivial bound for this is that this is less than N choose 4, but you know, suppose I am interested in pi 2 N or something like that, you know, maybe it is much, maybe it is much closer to order N, then N to the power 4 4. So, how does even one count that you will really depend on, you know, my ability to sort of. So, this this is a purely combinatorial problem, you know how do I draw those trapezoids to as to maximize this quantity.

So, this is worst-case, this is worst case over all arrangements of line segments. So, what can we say about this?

(Audio Not Clear. Refer Time: 28:56 to 29:04)

How can you assume that?

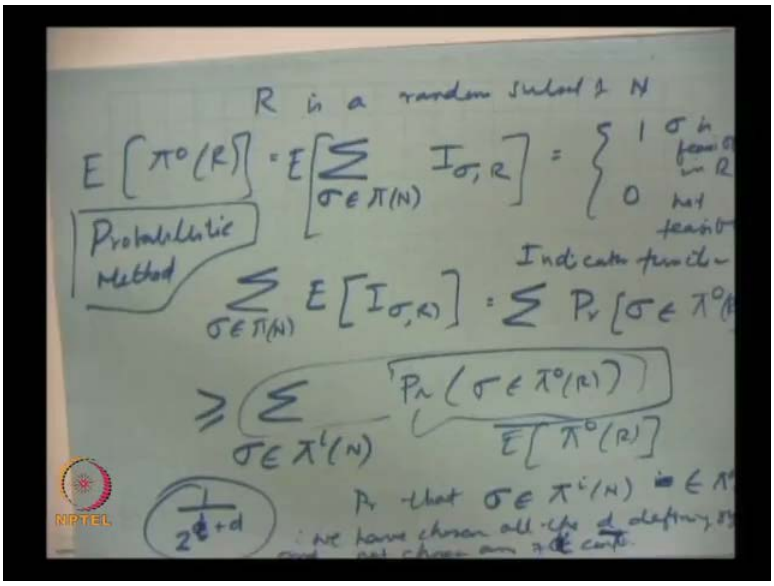
So, let me not even, you can think about it offline; let me give a simple way of proving this problem. So, there is a just say few lines of proof, let us let us go through that.

So, the claim is that pi i or I am written here a. So, pi i if I write it as I, pi i of N the size of this is bounded by big of 2 to the power i plus D expectation of pi 0 of R; R is a random sample R is a random sample of size N over 2, I have I deliberately pick the random sample of size N over 2. So, this quantity expectation when we pick a when we pick a random cycle of size N over 2.

Now, this is a very general bound, this does not talk… This this not just limited to trapezoidal maps, this is for you know other kinds of problem also if it is trapezoidal map and you pick a sample of sub random subset of size N over 2 what is the size of pi 0 i pi 0 of N over 2 order order N right N over 2 some some order N. So, this even this expectation actually, it it does not matter. So, it is actually, you know this is always less than, you know, some constant times N over 2; D is the number of number of objects defining a a configuration. So, this is the maximum number of objects defining a configuration any sequence.

So, mostly as I said, it is going to be constant so D is order 1. So, this quantity we go of 2 i plus D, if D is a constant and i is also a constant, then it follows that if I am interested in this pi 2 of N, in the case of trapezoidal maps, this whole thing is again order N, if you have to believe this claim. D is constant, because there is a maximum number of the maximum number of objects defining any configuration is 4; suppose I choose i to be equal to 2 that is a quantity of interested in pi 2 of N and pi… So, this is 2 the power some constant, which is constant, and then we are we argued that E expectation, it is not this expectation, in the worst-case also it is less than C times N over 2. So, this whole thing right hand side is big of N, if you are looking at pi 2 of N for trapezoidal maps. So, it is actually what we suspected that it is actually linear, and not 2 N square or N to the power 4 N to the power cube (( )) and why is that? Why? How can we claim this, you know from a random sample? So, that is what this proof is.

(Refer Slide Time: 32:28)



So, one way of, you know, describing pi 0 of R. So, R is a random sample, R is a random subset of I. So, what I have written I have written pi 0 of R is I sigma R and sigma is the set of all configurations of the N objects, and this I sigma of R equal to 1 or 0 depending on whether sigma is feasible in R or not feasible in R. So, if it is feasible in R, then we are saying that you know we will count that trapezoid if is not feasible, then we do not count the trapezoid that is it; it just a what is this is called indicator function.

Now, I take the expectation on both sides; see, this is for a specific subset R, but I am choosing R to be random subset right. So, I will take an expectation expectation I take expectation both sides. So, if we take then I can I can expand this using the linearity of expectation, and this is nothing but the probability that sigma belongs to… So, this if you go through the next couple of lines that prove what is it say.

So, E of pi 0 R is this just what I have written this is quantity of this, and then this quantity, so instead of summing over all pis of N if I sum only over pi i of N. So, since pi i pi of N is greater than that so I can say this is if I just restrict my summation to pi i of N. So, this will be greater than equal to that right. So, that is how that is all the proof is about and this is this is E of pi 0 of R this quantity this summation. So, pi a of N this is a right hand side is greater than or equal to pi a of N times, why did I write this one over 2 to the power a over t what is this quantity.

So, why this probability 2 to the power a plus d; every object is being chosen with probability half. So, so probability that sigma belongs to pi i N is sorry belongs to pi 0 of R means that we have chosen all the defining objects all the D defining objects, and not chosen any of the i conflicting objects. So, therefore, that probability is one over 2 to the power a plus d. The last line… So, so this quantity I am claiming is this quantity, why is is that? So, what is the probability that sigma belongs to pi 0 of R given that sigma belongs to pi i of N.

It means that - that particular configuration is feasibly R means that we have chosen all the D defining objects, and we have not chosen any of the i conflicting objects sorry a actually, I am using a here. So, well i or a sorry. So, this should be i sorry. So, that is that is one over 2 to the power i plus D right. All the i plus D has to be chased here right and. So, I will stop here today and the beauty of this proof is that this is a combinatorial bound that is completely independent of any kind of probability. I just use linearity of expectation to prove this bound, but this bound holds for any configuration of line segments. So, and this this method you know in literature is called the probabilistic method.