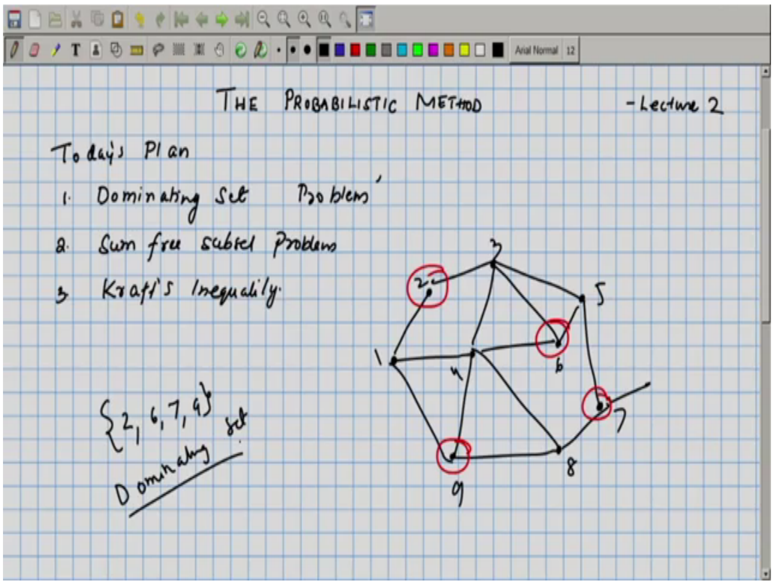
Randomized Algorithms
Prof. Benny George Kenkireth
Department of Computer Science & Engineering
Indian Institute of Technology, Guwahati

Lecture – 14
More Examples on Probabilistic Method

In today's lecture, we will be studying three problems. The first problem will be called dominating set problem that is a combinatorial problem from graph theory. So, we given a graph, we will define what is called as a dominating set. And we will use probabilistic method to bound the size of a dominating set. The second problem that we will look at is called the some free subset problem, we given a set of integers, and we need to construct a subset of these integers such that no two elements in that set adds to a third element in the set ok. The third problem is result from coding theory known as Kraft's inequality.

(Refer Slide Time: 01:20)

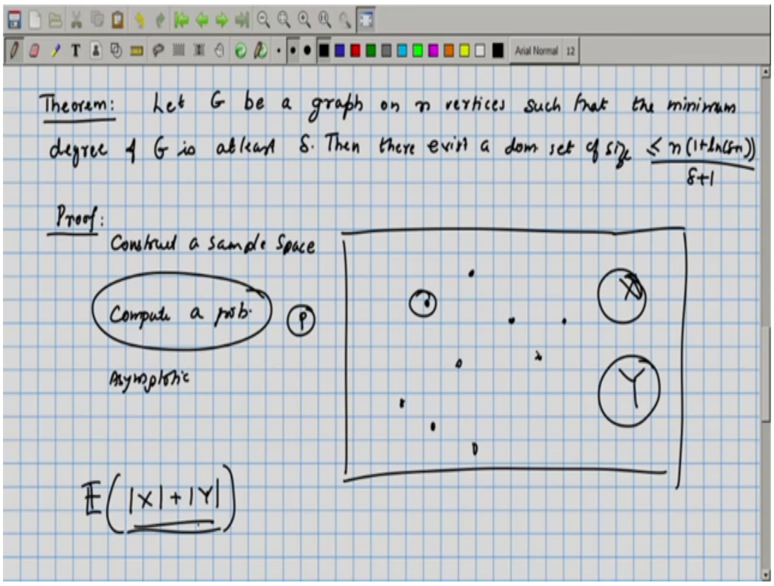


So, these this is the agenda for the day. The second being sum free subset problem, and third being Kraft's inequality ok; so, let us first look at dominating set problem. So, let us first define what understand what is known as a dominating set. If you given certain vertices and edges so given a graph ok so, we can consider a subset of vertices ok. And if we take these vertices, let me just number this 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10 sorry 9.

So, if you look at any particular vertex and ask this question, does it have a neighbour from the set that we have chosen. So, the requirement is for every other vertex, it should either be one of these red coloured vertices or it should be neighbour of one of these red colour, it should be a neighbour of one of these red coloured vertices such a set is called as a dominating set. So, here what we have chosen is not a dominating set, because if you look at vertex 7, this is not connected to any of the red vertices, so red vertices being the chosen vertices.

Vertex 1 is connected, 3 is connected, 4 is connected, 8 is connected, 5 is connected. So, except 7 everything else is we could have included let say vertex 7 itself. And then that collection of vertices that is 2, 6, 7, 9, this will be a dominating set ok. So, if we take all the vertices that surely is a dominating set ok, but we want to find a dominating set of the smallest size ok. What we will see today is a proof that a certain kind of graphs that will exists dominating sets of a certain size.

(Refer Slide Time: 04:31)



So, let us write down the theorem that we will prove ok. So, let G be a graph on n vertices such that the minimum degree of G is say at least delta, then there exist a dominating set of size less than n by 1 plus log delta plus 1 ok. So, how do we prove such a thing? Then there exist a dominating set of size less than n times 1 plus log delta plus 1 divided by delta plus 1, so that is a result that we will first prove. And the proof again will follow three steps of the probabilistic method first construct an appropriate sample space, and then we will compute a probability, and then we will do the asymptotics ok.

So, here we have to construct a dominating set, how do we construct the dominating set, the idea is simple. Every vertex, we will include or not include with the probability p ok. So, let us say if we look at one particular vertex, you toss a coin and the coin results coin toss results in a head, you include this into your dominating set, otherwise you discard ok.

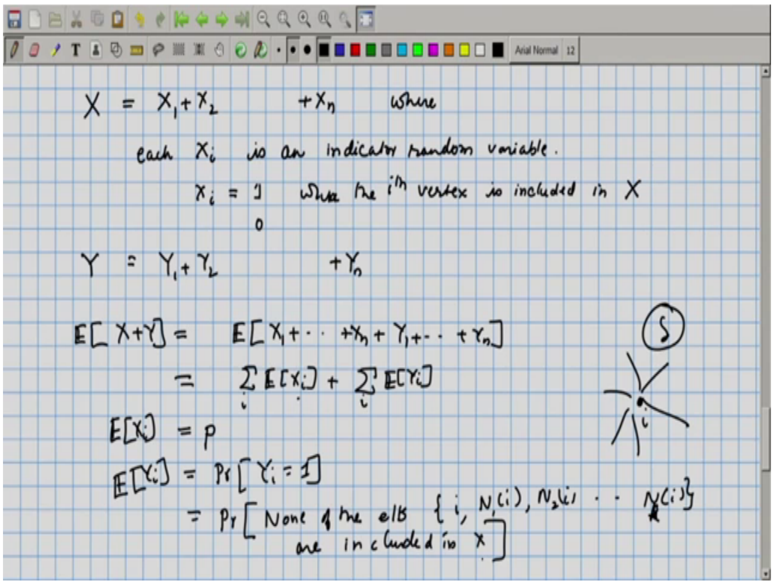
So, this way you will get a set X, which need not be a dominating set ok. But if it is not a dominating set, we will convert it into a dominating set by adding additional vertices ok. So, the additional vertices that we will add, we will call it is Y. So, Y will consist of all those vertices which are not X, and which does not have a direct neighbour which does not have a neighbour in X, so that is how you are going construct the dominating set.

So, the kind of sets that we will pick as a dominating sets are obtained in this particular manner. Toss a coin if the coin results in a head include that into our set, we will call that I mean all those vertices that we obtained, we will denote it by the set X and whatever additional vertices are to be added to this set to make it a dominating set that we will call as Y.

So, in the second step, what we will look at is what we are really interested in is the size of X plus Y ok. If we can show that the expected size of this is less than the quantity that we are interested in that is this particular number n times 1 plus log delta plus 1 by delta plus 1 if that is the case, then we know that there will surely exist one dominating set, which has size no more than that ok. So, this expectation computation will be our second step.

There is one important thing that I did not mention. I said toss a coin, but we could bias this coin instead of choosing a fair coin, we could assume that this coin has a bias p that is with probability p, we will include or exclude this particular point and this particular vertex and there are consideration from the dominating set ok. So, when we sample for the set X, we sample by using coins having a bias p. This p we will later on fix, so that it gives us the best possible bound ok, so that is the line of attack, let us look at the details.

(Refer Slide Time: 09:47)



Let us compute the expected size of X plus Y. So, X let us say size of X, let us call it as, we will just abuse a notation you just call it as X itself. So, the random variable X now denotes the number of elements that we include into the set X. So, this can be thought of as X 1 plus X 2 plus X n, where each X i is an indicator random variable.

We are going to see these kind of constructions very often, where we will express a random variable as sum of indicator random variables that will help us do quite a lot of computations effortlessly. But, the choice of the indicator random variable is very crucial, and what is the event it indicates that has to be chosen appropriately. Here X i denote, so X i is an indicator random variable, when it takes value 0 or 1 depending on certain condition it indicates a certain condition it is going to take the value 1, when the ith vertex is included in the set X ok, and 0 otherwise.

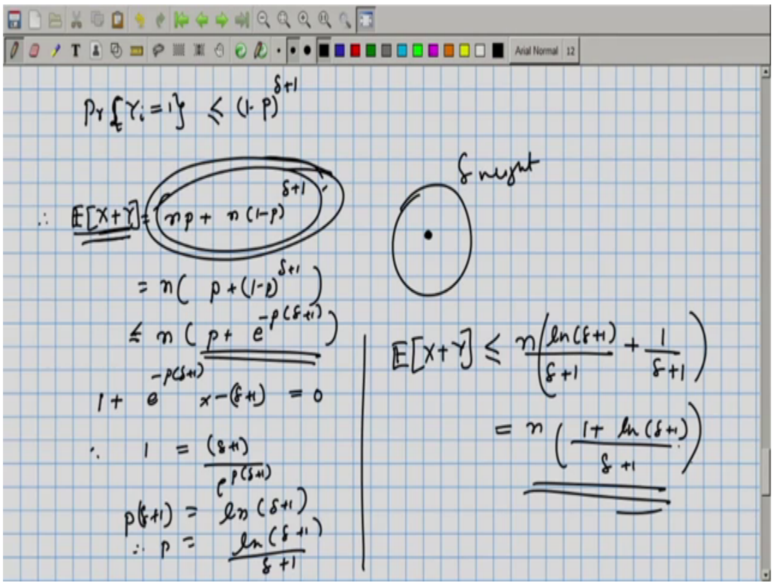
So, clearly the sum of the X i's gives the number of vertices that we included in the first phase of our sampling ok. And Y also we can let us say write it as Y 1 plus Y 2 plus Y n, where each Y i is either 0 or 1 depending on whether that ith vertex was included into our set Y ok, it is included Y i takes value 1, otherwise it takes the value 0 ok.

So, now what we are interested in is the expected value of X plus Y. And we know by linearity of expectation this is nothing but expectation of X 1 to X n plus Y 1 to Y n ok. So, we will apply linearity of expectation, and that gives us this expectation to be expectation of X i summed over all i plus expectation over Y i summed over all i ok.

And expectation of an indicator random variable is nothing but the probability that the random variable takes value one. So, this expectation of X i is equal to p, because we tossed a coin of bias p, and based on that we included it ok. And expectation of Y i will be what that will again be the probability that Y i equals 1. And when is Y i 1? Y i is 1 when, so let us look at the vertex the ith vertex, we know that it has at least delta neighbours, because the minimum degree of this graph is assumed to be delta ok.

So, only if i itself was not included, and none of the neighbours of i were included that is the only case in which Y i would have been included into the set Y ok. So, probability of this happening is same as the probability that none of the elements say in the set I mean if I write this as a set i, and neighbour 1 of i, neighbour 2 of i, and so on. So, let us say there are at least delta neighbour, and let us say if it is k neighbour N k of i are included in X only under these conditions will Y i be 1.

(Refer Slide Time: 14:47)



Therefore, we can say that this probability that Y i equal to 1 is less than 1 minus p raised to delta plus 1 that is because this element, and the delta neighbours none of them should be included. And since the choice is being done, independent of the other choices this probability is less than 1 minus p raised to delta plus 1 ok.

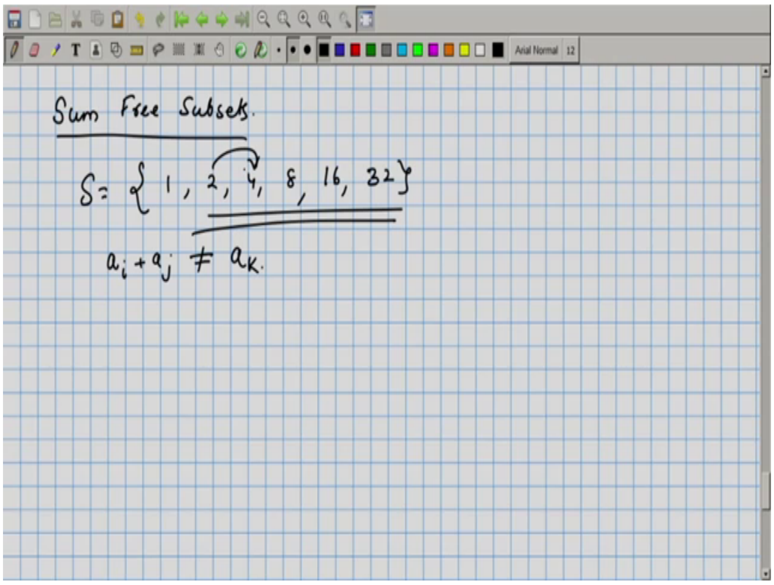
So, therefore expectation of X plus Y is equal to n times p plus n times 1 minus p raised to delta plus 1 ok. What we are interested in is bound in this expectation, so what is the value of the first question, we can ask ourselves is: what is the best value of p to make this quantity as small as possible. You can see that this is a function of p ok, so this is equal to n times p plus 1 minus p raised to delta plus 1 or we can just rewrite this.

We can write this as this is going to be less than or equal to n times p plus 1 minus p is less than e raised to minus p, so this quantity is anyway less than n times p plus e raised to minus p times delta plus 1 ok. So, this is just to simplify the calculations, we need to minimize this quantity by choosing p appropriately. So, we can just differentiate it with respect to p, and we will get and equate it to 0. So, n is anyway going to be a constant, we would not bother about it.

So, 1 plus differential of e raised to minus p delta plus 1 is e raised to minus p delta plus 1 times minus delta plus 1, this is going to be 0 at the minimum value. Therefore, 1 is equal to delta plus 1 by e raised to p delta plus 1, which simplifies to p times delta plus 1 is equal to log delta plus 1, therefore p can be chosen to be log delta plus 1 by delta plus 1. So, this is the best value for this is value which minimizes this expression.

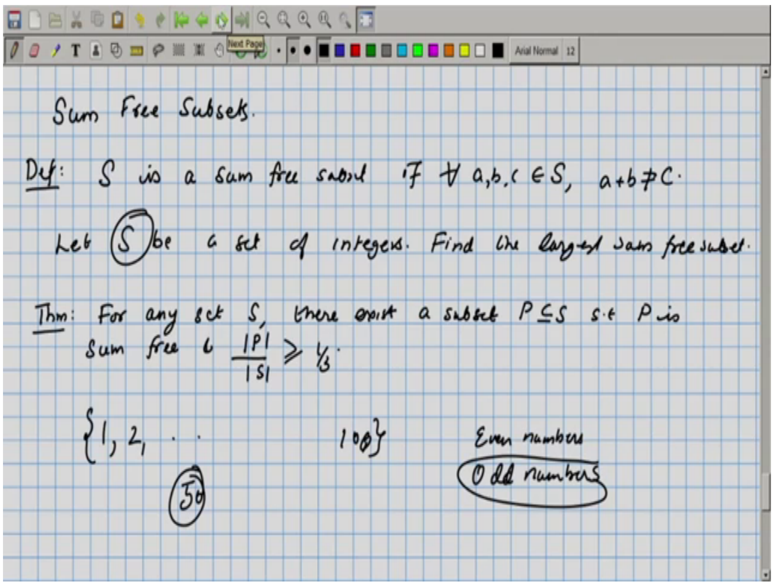
If you plug in that value, we will get expectation of X plus Y is surely less than n times log delta plus 1 divided by delta plus 1 plus e raised to p delta plus 1 is equal to 1 by delta plus 1. So, from this we know that e raised to 1 by e raised to p delta plus 1 is 1 by delta plus 1 that is equal to n times 1 plus log delta plus 1 divided by delta plus 1. So, since the expectation is less than this quantity, there will surely exist a sampling which contains no more than these many vertices ok. In other words, there exist dominating set of size at most this much, so that concludes the proof.

(Refer Slide Time: 18:39)



We will see our next problem. So, we will so the problems is that of sum free subsets. So, let us define what is the free subset. If we look at let us say 1, 2, 4, 8, 16, and 32. If you take any two of them and add, you will not get another element in the set ok. So, S has this property that if a i plus a j ok, here i is not equal to j will not be equal to a k ok. So, this is not a sum free subset, because we could choose i and j to be equal, and then we will get and 2 plus 2 equals 4.

(Refer Slide Time: 19:49)

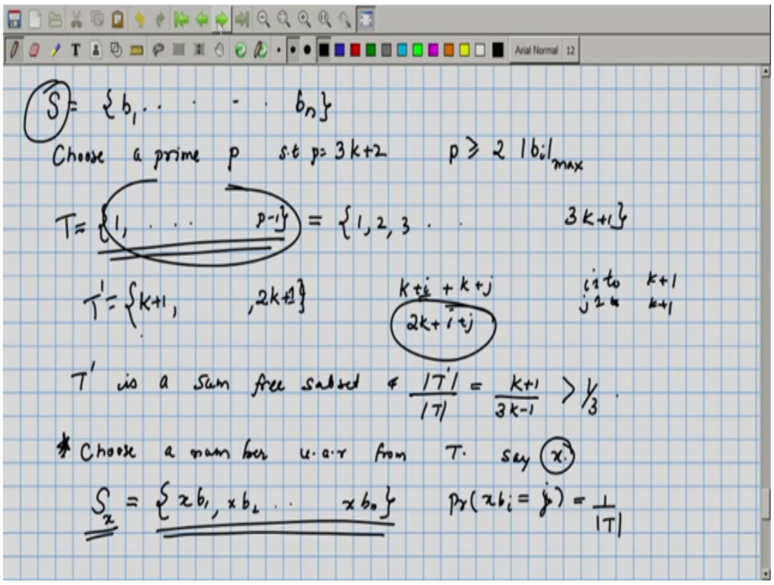


So, a sum free subset is a set of numbers such that the sum of any two elements in the set is not present in the set ok. So, definition S is a sum free subset, if for all a, b, c belonging to S, a plus b is not equal to C ok. Now, we given a the problem that we will look at is the following. We given a set of integers, so let S be a set of integers find the largest sum free subset ok, we are interested in interested in bounding the size of the largest sum free subject. We want to say that there exist a large sum free subset for any set S ok.

So, the theorem that we will prove is for any set S, there exists a subset P of S such that P is sum free, and size of P divided by size of S is greater than or equal to one-third ok. So, you can construct a sum free subset of size at least one by one by third of the original size. Let us just take a couple of examples. So, if you take the set of numbers 1 to 100 ok, if you take the collection of all even numbers that is not a sum free subset. But, if you take the collection of all odd numbers, it will always be a sum free subset, because for any two odd numbers if you add them, the result is going to be an even number, and that is not going to be present in this.

So, there exist a sum free subset of size 50. One can ask is there something of size more than 50 ok, so we in this case we were lucky enough to construct a sum free subset, which is half the size of the original one. The theorem states that in any case, you will always no matter what the set S is you can construct a sum free subset of size at least one-third of the original size. So, we will see proof via probabilistic method.

(Refer Slide Time: 22:47)



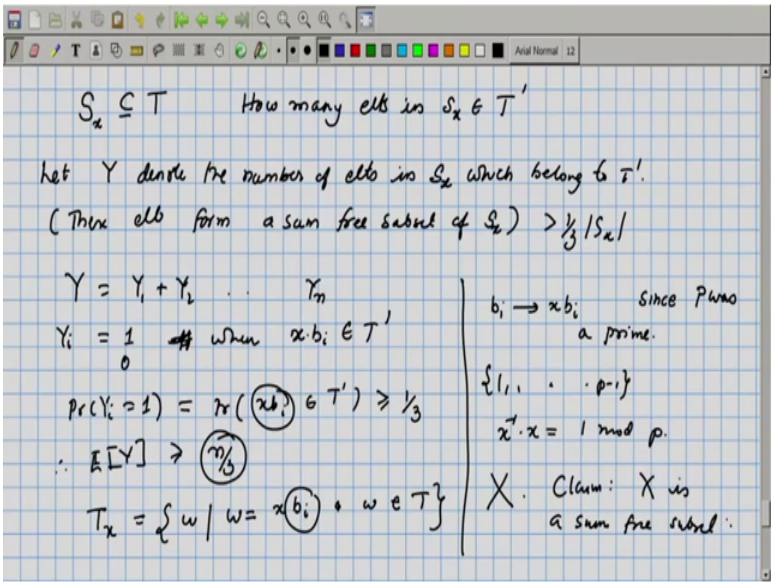
So, let us call this set S as b 1 to b n ok. Now, you will choose a prime p such that p is of this form 3 k plus 2 ok. One can show that there are infinitely many primes of this kind ok. So, additionally this prime p that we choose should have the property that p should be greater than 2 times to mod b max. So, look at all the b i's look at its absolute value, whatever is the maximum twice of that should be smaller than the prime that we choose ok, we can of course choose one such, because they are all infinitely many primes of the form 3 k plus 2.

So, we are going to this randomly sampled the elements of the set S to construct our sum free subset ok, but this sampling has a clever sampling ok, we will do it in phases. First let us look at another set, which is basically numbers from 1 to p minus 1 ok. So, this is exactly same as the set 1, 2, 3 up to 3 k plus 1 ok. Now, in this set if you look at the numbers from k plus 1 to 2 k plus 1 ok, they form a sum free subset of this set. If you call this set as T, and this is T prime, T prime you can think of the middle third of T, they form a sum free subset. Because, you take any two elements in this, they are of the form k plus i plus k plus j, where i can vary from 1 to k plus 1, and j also can way from vary from 1 to k plus 1.

So, if you add them, we will get 2 k plus i plus j ok. And the largest element so this is at least this much. And i and j being at least one, this number is going to be at least 2 k plus 2. So, any two element in this set is going to result in a sum, which is greater than every element in T prime. So, T prime is a sum free subset. And size of T prime divided by size of T is equal to they are k plus 1 elements there divided by 3 k minus 1 right 1 to p minus 1, so this is greater than one-third ok. Now, how do we use this fact to construct the sum free subset of S, what we will do is the following, choose a number uniformly at random from T. So, let us call it as x ok, we will take this x, and multiply every element of S with x ok. So, we will get a set which we will call as S x, so S x is going to be x b 1, x b 2, x b n ok, what have we achieved by this construction in this operation ok.

So, initially b 1, b 2, b n where some integers, but here we have made sure that this set S x consists of random elements. Since we are choosing x uniformly at random, x b i for any b i is going to be distributed uniformly at random over this particular set 1 to p minus 1 ok. So, you can say the probability that x b i is equal to let us say i or j i mean for j belonging to 1 to p minus 1. This is going to be equal to 1 by size of T here. When we choose an element uniformly at random from the set and multiply this, we are essentially choosing another element uniformly at random from T. Now, how why did we choose this prime p well that helps us in a certain way that we that will be made clear.

(Refer Slide Time: 28:05)



Now if you look at this set S, S is a subset of T, we can ask this question or S x, how many elements in S x belong to T prime. So, T prime was this middle third x was chosen uniformly at random, this number the number of elements of S x belonging to that will also be uniformly distributed I mean, they will be I mean they will be random. So, let us look at each element the probability and look at so we will compute the following quantity. So, let us say Y denote the number of elements in S x which belong to T prime ok, what does Y indicate I mean why are we interested in Y well this basically tells I mean, any element in I mean any I mean if you take these elements which belong to T prime, they are going to form a sum free subset of S x ok.

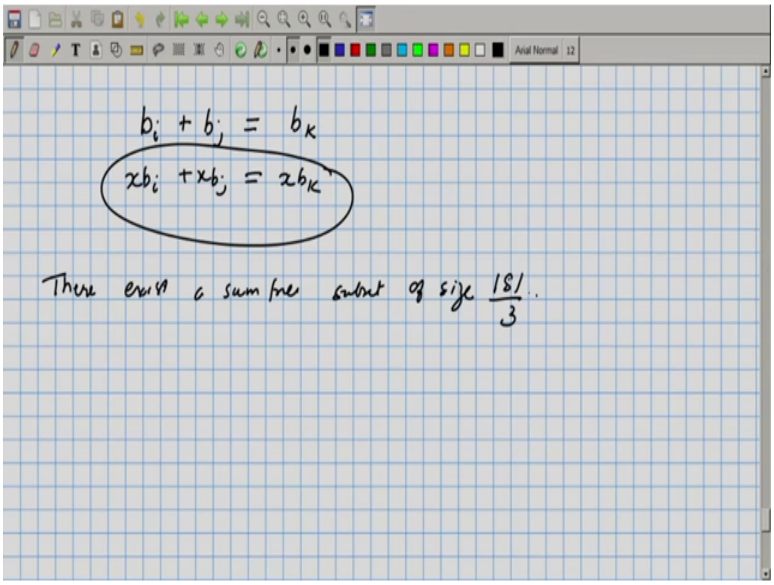
These elements form a sum free subset of S x ok, and their size is going to be greater than one-third of size S x why so, so we can again look at this the element we can write Y is equal to Y 1 plus Y 2 plus Y n, where n is a number of elements in S. Y i equals 1, if and only if I mean so 1 n it takes value 0, and so Y takes value 1, when x times b i belongs to T prime ok.

And the probability of that happening we know is going to be so for each Y i probability that Y i equals 1 is going to be less than or equal to probability that x b i belongs to T prime. And that since x b i is uniformly x b i can be thought of as chosen uniformly at random from the set from this set t, we know that this probability is going to be certainly sorry I should not write less than here. So, this probability is equal and this has to be greater than or equal to 1 by 3 ok.

Therefore, expectation of Y is going to be greater than n by 3 ok. So, this tells us that there exists a subset of S x of size at least n by 3 which is sum free, but we were interested in the subset of S, and not S x ok, but what we can do is the following. So, let us say we call that subset by T x, so T x if we say are all the elements w such that w is equal to x times b i, and w belongs to T ok, all those elements which belongs to the middle third ok.

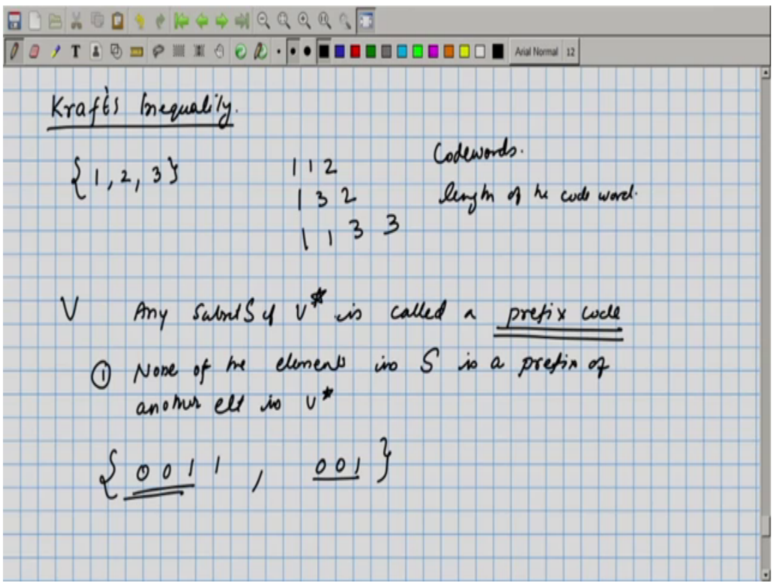
If you take that as T x, what we showed is there is a T x of size greater than n by 3. And if we look at the pre majors or the b i's corresponding to that ok, so if this map from b i to x times b i, this is an invertible map since your p was a prime. So, if you take the element 1 to p minus 1, there is an element x inverse such that x times x inverse is equal to 1 mod p ok. So, for every element in this set T x, if you just multiply it with that x inverse, you can obtain a set S ok.

(Refer Slide Time: 34:05)



Further, if you can obtain a subset of S, let us call that as X, so that subset if you call it as X, we can claim X is a sum free subset. So, why is this so if not there are two elements in an in X, let us call it as b i and b j. So, b i plus b j is equal to some b k ok, but if this was the case, then if you take x b i plus x b j that is going to be equal to x b k ok. But, x b i, x b j, and x b k, they are all elements in this particular set T x ok. And they we assumed is going to be sum free. So, this basically tells us that there exist a sum free subset of size S by 3, and greater than or equal to size S by 3.

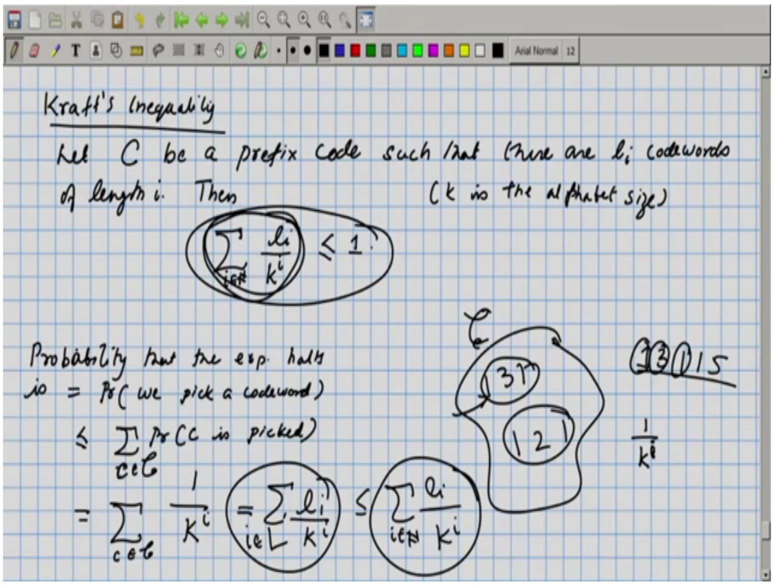
(Refer Slide Time: 35:04)



Now, we will look our third problem, which is a problem from coding theory is called this is an inequality called as Kraft's inequality ok. So, let us understand what this inequality is about say let us take a set a finite set of three elements ok. We are going to construct strings using these elements by strings I mean you can look at 1 1 2. So, this is a string of length three 1 3 2, 1 1 3 3 ok, these things we will call as code words ok. And the number of letters that appear in a codeword is will be called as a length of the code word ok.

If you look at any subset of v star, that will be called a code that will be called a prefix code, if certain conditions are satisfied. The condition is none of the elements in S is a prefix of another element in v star. For example, if you had taken 0 0 1 1, and 0 0 1 ok, this is not a prefix code, because this is present as a prefix of another codeword. So, this collection will be called as a code, and individual strings will be called as code words. A collection of strings such that no codeword is a prefix of another codeword will be called as a prefix as a prefix code. Kraft's inequality tells us about some condition that should be satisfied by all prefix codes.

(Refer Slide Time: 37:36)



So, let us write down the condition ok. So, let C be a prefix code such that there are l i code words of length i. Then summation over l i k to the i is less than or equal to 1, where k is the alphabet size ok. So, prefix code over a particular alphabet of size k ok.

So, this is called as Kraft's inequality; slight variant from the kind of probabilistic method applications that we have seen so far.

In the sense, here we are not going to let us say do any complicated asymptotics, we are just going to say at a certain probability, we will describe a certain event. And this inequality will come almost for free, because the left hand side will indicate the probability of a certain event ok. And since, it is a probability of a certain event that has to be always less than one. So, you can try and imagine, some event in an appropriately constructed probabilistic space such that its probability is exactly this. If this happens to be the probability of some event, we can say that since it is a probability it has to be less than 1 ok.

So, let us construct our experiment, and then the event will also be clear ok. So, what we do is a simple experiment. So, randomly keep on picking elements from the set 1 to k. So, your alphabet let us say it was elements from 1 to k keep on randomly picking symbols from 1 to k ok. So, may be when I first pick I get it 1 and 3 1 1 and so on ok. I keep on doing this keep on writing the symbol that I picked, and I stop, only when I get a code word ok, I could keep on going for ever. So, there is a an already fixed collection of code words which is given to us let us call let us C ok.

And so let us say if C contain 1 3 1, and 1 2 1 ok. I pick 1, and let us say again 1, and then this process is never going to end. This process we will terminate, this process of continuously picking symbols on the alphabet we will terminate only when I hit one of these code words, otherwise the process we will continue forever. The event that I am interested in is that I will hit one of code words. So, what is the probability that I will stop picking let us ok, it is that is going to be this.

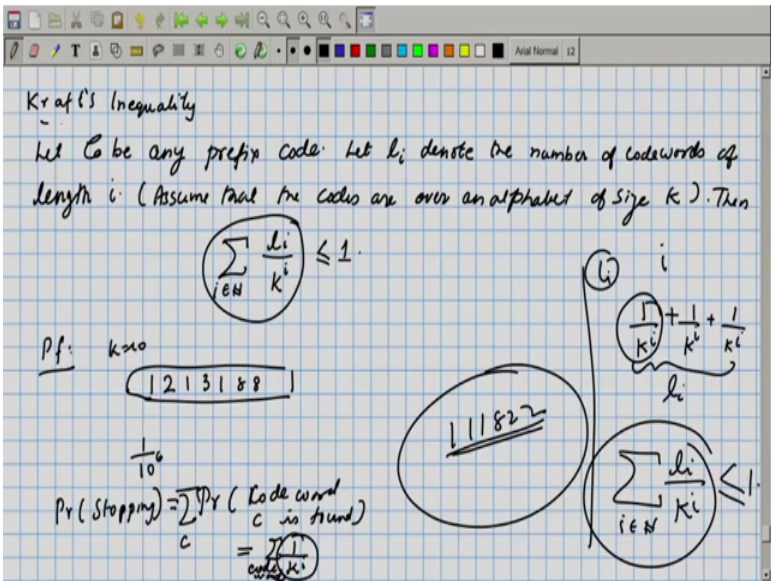
So, you can see that since its prefix code for each element so if you pick 1 3 1, then of course it precludes the possibility of picking any other code word ok. For a particular code word what is the probability that, so we will stop only if so probability of stopping that the experiment halts is equal to probability that we pick a code word ok, so that is going to be less than sum over each code word probability that. So, C belonging to the collection of code words sum over each code word probability that C is picked.

Now, since each element I mean the way we are constructing this is by sampling from 1 to k, the probability that a particular code word let us say if we had the code word 1 3 1 1

5, this is picked only if the first pick is 1, and the second pick is 3, the third pick is 1 and so on. So, when did I when did when did we stop, we stopped only if we discover a code word as the prefix that can happen only if the first pick is let us say if one, the second pick is 3, the third pick is 1, and so on, so that happens with probability exactly 1 by k raised to i ok, where i is the length of the code word ok.

So, this probability is equal to sum over all code words c belonging to C 1 by k to the power i ok. So, this is going to be nothing but summation l i by k to the power i i. So, i varying overall length, so I will just write this as set l, where l is the set of all possible length ok. So, this quantity l i by k i ok, so here I should have written then summation l i i belonging to natural numbers. So, this is less than summation over i belonging to natural numbers l i by k to the power i. So, since this is the probability ok. So, this is a since this quantity is a probability that is going to be less than 1 ok, so that is the proof.

(Refer Slide Time: 44:02)



So, let us see Kraft's inequality. So, let C be any prefix code let l i denote the number of code words of length i. Assume that the codes are over an alphabet of size k. Then summation i belonging to natural numbers l i divided by k to the power i this is going to be less than or equal to 1 ok. The proof will again be by probabilistic method, but the proof is going to be slightly different from the earlier probabilistic method proofs that we have seen. So, here we will just compute a certain probability, we will we will describe a certain event, and we will say that the probability of that event is at most this.

So, if the probability of that event is this, and that being a probability has to be less than 1 ok. So, what is the event that we are interested in what is the experiment ok. So, let us say we keep on randomly picking numbers from 1 to k ok. So, let us say if k equals 10, then I will just pick numbers of the following from 1 2 1 3 1. And this process I will keep on continuing this till I get a code word somewhere during my pick.

So, if I have obtained a code word at that point I stop, otherwise I will continue forever ok. Now, let us look at the probability that we will stop at sometime, we will stop only if we find a code word ok. For example, if 1 1 1 8 2 2 was there, we will stop only if we get this particular string. And the probability of getting this particular string while we randomly pick a sequence is 1 by 10 to the power so here k, if it was 10 into it would have been 1 by 10 to the power 6 that is the only way we can get this.

And since it is prefix free, we can say that I mean if you get 1, you will not get another ok, because otherwise one string would be prefix to the other ok. So, probability of stopping is equal to probability that code word C is found, and they summed over all C ok. So, this is nothing but for each C, the probability is 1 by k to the power i, where i is the length of the code word, this summed over all code words ok.

So, each code word of length i gets picked with probability 1 by k to the power i. If any one of those are being picked we will stop, otherwise we will continue. And that probability is the sum 1 by so if there are l i code words of length i, one of them gets picked with 1 by k to the i plus 1 by k to the power i by 1 by k to the power i, there are i there are l i terms ok. And therefore, the total probability is just summation over i belong into naturals l i by k to the power i, and since this is the probability. Of course, it is going to be less than 1 that concludes the proof of Kraft's inequality.

Thank you.