Probability for Computer Science

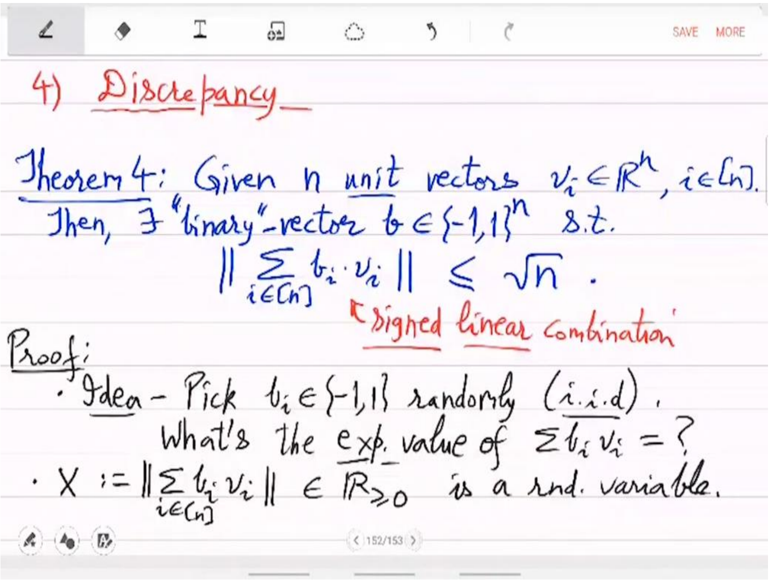
Prof. Nitin Saxena
Department of Computer Science and Engineering
Indian Institute of Technology - Kanpur

Module - 8
Lecture - 29
Extremal Set Families

So, last time we were doing probabilistic methods. So, we will be doing many examples. First we showed the example of Ramsey number, which is also you are colouring the graph, the edges red or blue. And then, you are interested in monochromatic complete subgraphs. This you can also see as the presence of a clique or an independent set. You can think of complete graph on red as clique and blue as absent edges, which means that it is an independent set.

Second example we did was about the existence and also finding a large cut in a graph, a subset so that the edges going out are maximised. Then we did sum-free subset. So, given a set of numbers, finding a subset which is a devoid of any sum. So, A + B should not be equal to C; A B C in the same subset. And then, we have started discrepancy.

(Refer Slide Time: 01:37)



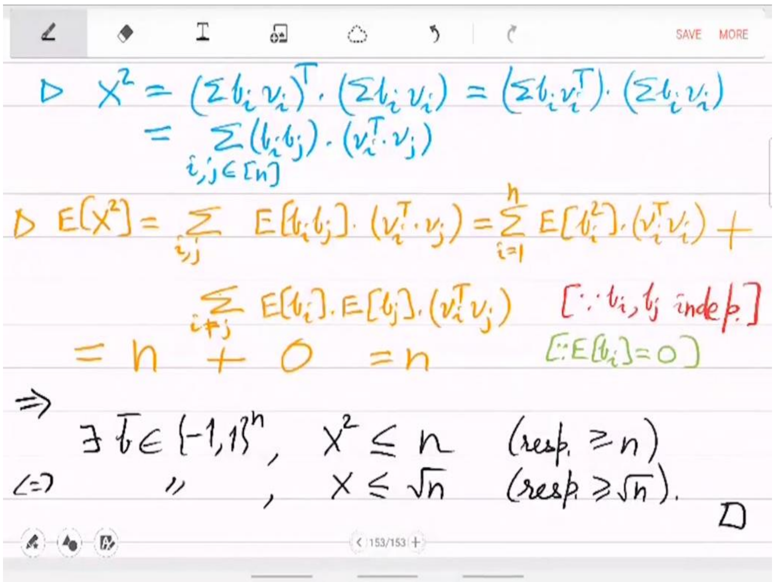
So, here, you are given n unit vectors. You do not have any information about v i's, except that they are in ambient space n, real coordinates. Now, you want to take a signed sum, this Sigma b i v i, so that the norm is minimised. And we will show how to find a signed sum where the norm is smaller than square root n. So, note that the signed sum will have norm;

you expect it to have norm around n, because v i's are unit vectors; but this is far better than that; this is giving you square root n.

Idea here is, as should be clear by now that we will pick b i's randomly. So, you pick b i to be -1 +1, by a coin toss randomly. In fact, you do it in an IID fashion, which is independent and identically distributed. So, you flip basically n distinct coins independently. So, what is the expected value of Sigma b i v i? What do you expect this to be? So, let us do that calculation. So, define a random variable X as the square root of this; let me first do this; define it to be the norm of b i v i.

So, this is what we are interested in. So, what is the expectation of this? Remember that this is a real-valued, positive real-valued in fact. So, this is a non-negative real-valued random variable. And now, we will calculate its expectation.

(Refer Slide Time: 04:28)



So, X square is, essentially it is Sigma b i v i transpose multiplied by itself. So, this is the inner product, in other words. Advantage is, you can expand it out. So, this is equal to Sigma b i b j times v i transpose v j. And let us instead calculate the expectation of this, expectation of X square. So, that is, by linearity of expectation, you get b i b j; and v i transpose v j is fixed; fixed value. So, that goes out.

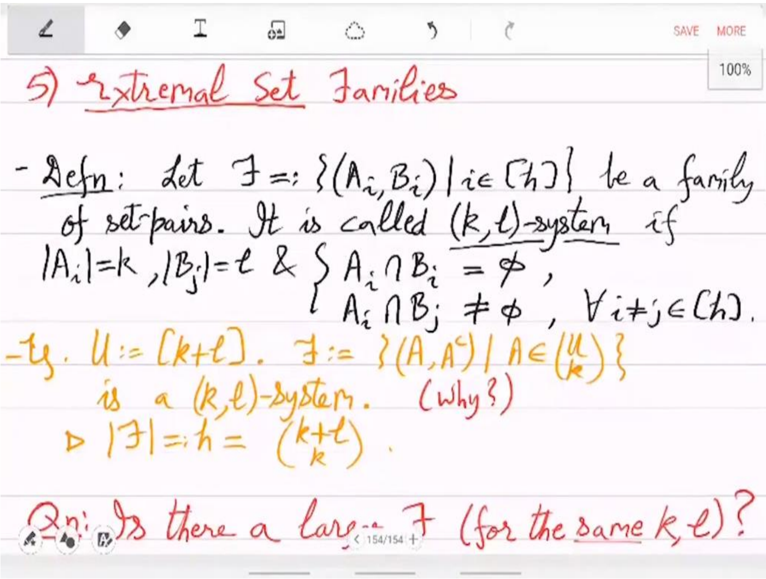
Remember, b i b j is what you are picking randomly in an IID fashion. So, if i and j are different, then, because of independence, this will factor; otherwise, you are looking at expectation of a square. So, let us do that. So, expectation of b i square times this plus expectation of b i b j times v i transpose times v j. So, this is equal to expectation of v i square; so, v i square is 1. So, that is 1.

And you are looking at v i transpose v i, which is actually norm of v i, which is also 1. So, you get n plus; what is expectation of b i? Expectation of b i is 0, -1 +1. So, you get n. So, here we used independence. And notice that actually 2-wise independence would have been enough for this calculation. And the other thing we are using is that expectation of b i is 0. So, that gives you expectation of X square to be n, which means that there indeed exists this choice b 1 to b n such that X square is less than equal to n.

Since the expectation of the random variable X square is n, so, there has to be a choice such that its value is less than equal to n, respectively greater than equal to n; nothing special about that. So, which means, notice that X is by definition, it is non-negative. So, this is if and only if there exists b bar such that X is less than equal to square root n, respectively greater than equal to square root n.

So, we just wanted to show that there is a choice of b i's, such that this X is at most square root n, and that is true. That is what expectation has shown. So, this is a useful calculation. It actually appears in many places; this expectation of square computation and use of 2-wise independence.

(Refer Slide Time: 09:11)



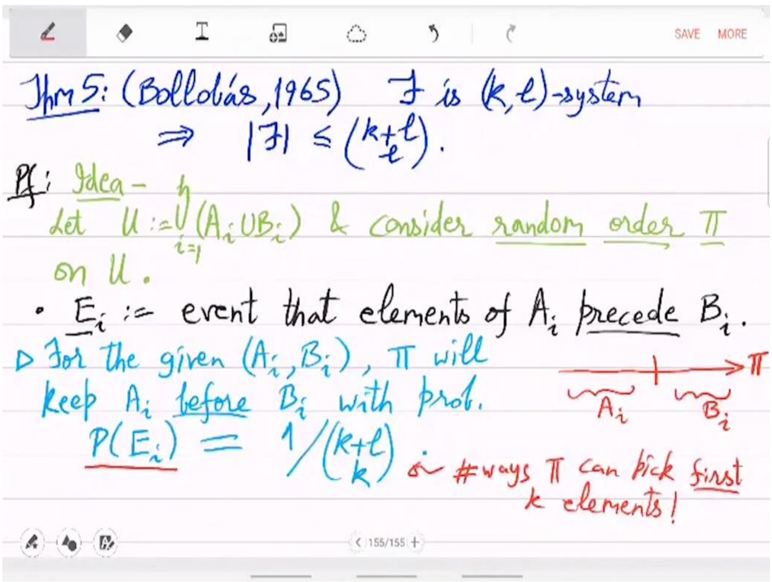
So, we have seen 4 examples of probabilistic methods. Now, we will work out the fifth one, which will be slightly more complicated. So, this is about extremal set families. So, what is that? So, define a family of set-pairs, let us say h pairs. It is called (k, l)-system if A i's are of size k; B i's of size l. Let me call this B j. And if you look at the intersection of A i with the other set B i in the same pair, they are disjoint for all i.

And if you look at the intersection with something else, then it is not, for all i different from j. So, again, A's are of the same size, B's of the same size, which is respectively k and l. In the same pair, these 2 sets are disjoint; but across pairs, they overlap. That is called a (k, l)-system of sets. So, let us see an example first. So, let the universe be first k + l numbers. Then, if you look at the family where first is a subset of u, and the other is its complement, and take A to be a subset of u of size k; thus, A compliment has size l, because universe has k plus elements; so, A takes k and A compliment takes l.

So, you can see that in a single pair, A compliment are disjoint, while if you go across pairs, then they will overlap. So, you can see as an exercise that this is a (k, l)-system. So, you can check this. And what is the size of this (k, l)-system h? So, this size is k + l choose k. That is the number of subset A that you can have of the universe. So, the question to ask is, is there a larger F for a fixed k, l?

So, given k and l, can you find k, l’s system where the number of set-pairs is more than k + l choose k? That is the question. Now, a priori there is simply no intuition whether the answer is yes or no. You can try at this point, either direction.

(Refer Slide Time: 14:19)



What we will show, that is theorem 5. We have done till theorem 4. So, theorem 5 is due to Bollobas, saying that F is (k, l)-system. Then, if F is (k, l)-system, then previous example is the best. That is, F is smaller than k + l choose l. So, any (k, l)-system of set-pairs cannot have size more than what we just saw; k + l choose k is the best. This upper bound seems stunning, but the proof is really simple by probabilistic method.

So, what you have to do is kind of take inspiration from this example. So, in this example, the universe was ordered; you had this 1 to k + l. So, kind of map every example to this example via ordering. So, the idea is, define the universe to be A i union B i for all i, 1 to h, and consider a random order pi on u. So, randomly map this universe to kind of what you saw before. And what is the contradiction or what is the argument now? How does it help?

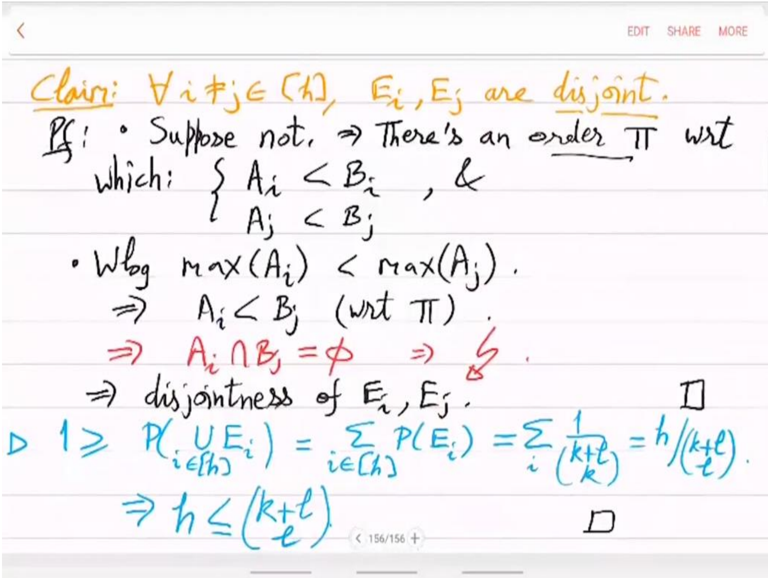
So, for that, consider the event E i, event that elements of A i received B i. In this order pi, what is the chance that in a set pair, in the ith set pair, A i has elements all smaller than all the elements of B i. So, there is a marker; A i is completely on the left and B i is completely on the right, with respect to order pi. So, this is event E i. What is the probability of this happening?

Well, so, assuming that we have fixed pi or you have chosen pi, A i is essentially k elements that you want to pick. You will pick these k elements. And B i are l elements; you have to first remember that A i, B i are also given. A i, B i is fixed. So, let me write that. So, for the given A i, B i, pi will keep A i before B i with probability. So, basically, you just focus on these k + l elements.

And what is the chance that pi will pick the first k before the remaining l? So, that is just the probability of choosing a k subset out of k + l. So, it is 1 over k + l choose l; k + l choose k also you can say. So, this is the number of ways the first k can be chosen. And out of this, A i is the only correct one. So, it is actually one out of all these possibilities. That is the number of first k; number of ways pi can pick first k elements.

And out of this, only one is correct, which is A i. So, this is 1 over k + l choose k, which is equal to 1 over k + l choose l. So, that is the probability of E i, event. And what can you say about these events? So, E 1 and E 2, or in general, E i and E j, for different i j, are they dependent? Are they independent? Are the disjoint? What can you say about that? So, I claim that E i and E j are disjoint events.

(Refer Slide Time: 20:39)



So, this may not be clear immediately, but once you think about the definition of disjointness, you will understand. So, suppose not, suppose they overlap, there is some common case. So, what is the common case? Suppose there is a order pi; so, there is an order pi under which or with respect to which A i is all behind B i and A j is also behind B j. So, first is event E i; the second is event E j; and there is a common order pi which is achieving both the things.

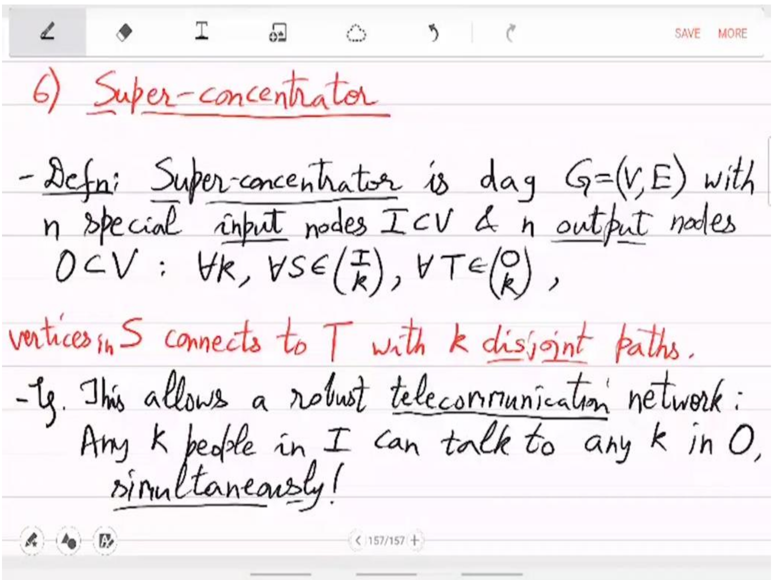
Now, is this possible? Well, let us assume without loss of generality that max of A i; so, compare max of A i with max of A j. So, suppose max of A i is less than max of A j; but max of A j, you know is less than every element in B j. So, A i goes completely behind B j with respect to pi, which means what? In particular, it means that A i and B j cannot intersect, which is a contradiction; which means that E i, E j are disjoint; they do not have a common pi.

So, in some sense, this E 1 to E h, they are giving you a partition. So, E i is saying, is the event that A i precedes B i. And so, if A 1 precedes B 1, then A 2 cannot precede B 2. This is what we have learnt. So, what you can deduce from this is that 1 is; if you look at the union of E i's, union of these events, then, that is actually sum of the events. So, you are actually getting a partition of union of E i, and that of course is less than or equal to 1.

Probabilities are always less than equal to 1; so, we have that the sum is less than equal to 1. But what is this sum? What is probability of E i? So, that we have calculated. That was 1 over k + l choose k. So, you have calculated h over k + l choose l, which implies that h is less than equal to the binomial number or the binomial coefficient. So, we have shown that by this ordering, by picking a random order, you actually get an upper bound on how many A i, B i's are there.

So, this is a very short proof, relatively simple, but it is proving this statement which had nothing to do with probability or ordering, completely devoid of any character except this set family properties. So, with this, let me move to the final example of probabilistic method.

(Refer Slide Time: 25:54)



So, sixth example is super concentrators. So, these are again, kind of extremal graphs; these are extremal situations. It is a graph which is highly connected. And we will make it precise. And then, using probabilistic method, we will actually give a way to construct them. So, let us first see the definition. So, super concentrator is a directed acyclic graph, G. So, basically, the edges are directed, and there is no cycle.

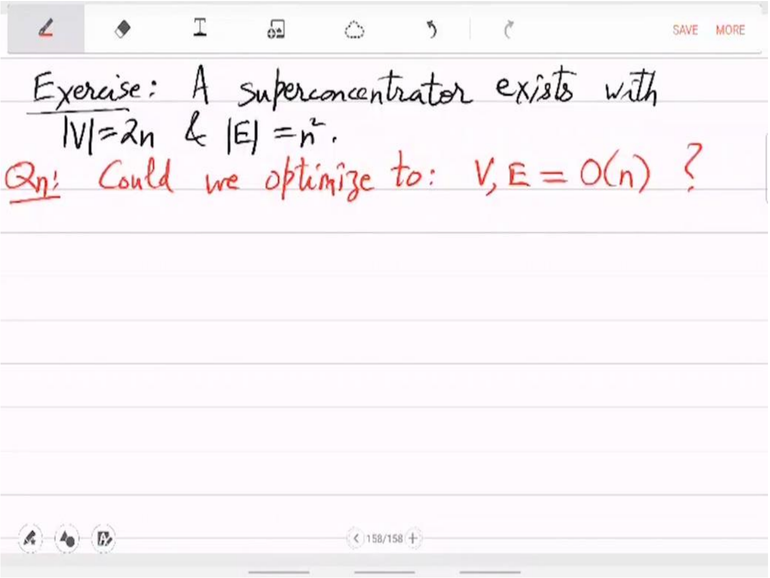
It is a directed acyclic graph with 2 designated subset of vertices. One we will call input, the other we will call output. So, with n special input nodes I, and n output nodes O, such that; so, here comes the heavy connectivity between input nodes and output nodes. So, no matter what subset you pick of I, respectively that of O, you will have disjoint paths between them or across them.

So, for all k, for all S, I to k and for all T; sorry, S is a k subset of I, I choose k and T is a k subset of output nodes; so, O choose k. For all ST, S connects to T with k disjoint paths. So, the vertices S, they connect to vertices T. So, k on the left, k on the right, you can think; but via k disjoint paths. So, what is the importance of disjoint, k disjoint paths? The thing is that, if you think of like a physically wired telecommunications network, then, for people in the set S to talk to people in the set T, you want really disjoint connections.

If the connections overlap, then the calls will overlap. So, that really is the motivation of this definition. So, example: This allows a robust telecommunications network. So, any k people in I can talk to any k in O simultaneously. This is the importance. So, any k people in input nodes can talk to any k in output nodes. So, you can think of them as the communities; and simultaneous connection is there; simultaneously they can talk.

So, these are the wires. These wires do not overlap. Now, do these things even exist. So, actually, existence is not a problem, because a complete graph. And then, these connections will be there. So, these things trivially exist. So, let us note that down.

(Refer Slide Time: 31:06)



So, a super concentrator exists with vertices 2n, and edges n square. So, you can look at the bipartite graph; n vertices on left, n vertices on the right; all the edges are there. So, then, trivially you have, any subset on the left can talk to any subset on the right, via this matching. So, that is actually trivial. n square is there. So, what is the question then? The question is, could you reduce the number of edges? Could you really get a linear sized graph?

So, could we optimise? let me say. So, that vertices and edges, both are linear in n. You want the vertices and the edges to be roughly of the same size; and hence, linear construction. And that existence is now not clear.