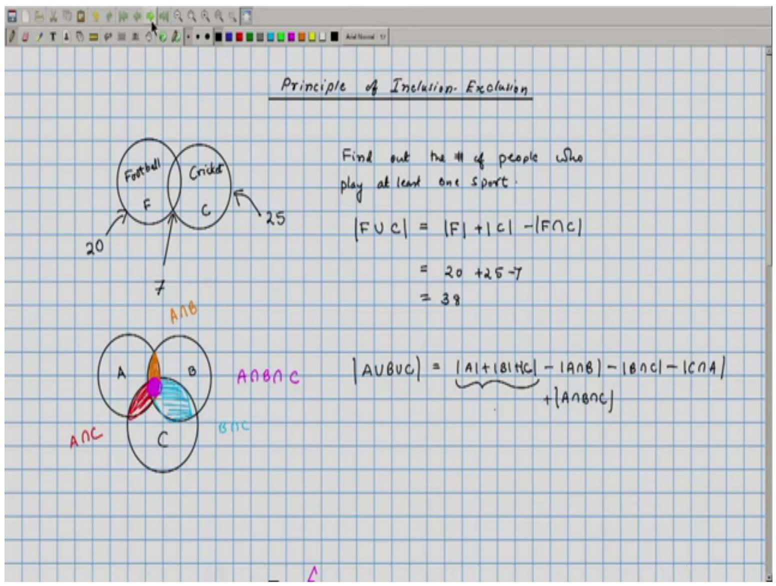
Discrete Mathematics
Professor. Benny George
Department of Computer Science & Engineering,
Indian Institute of Technology, Guwahati.
Lecture 32
Principle of Inclusion Exclusion

In this lecture, we will learn about the Principle of Inclusion Exclusion.

(Refer Slide Time: 00:32)



So, let us start by the following example. Let us say, this is a set of students who play football, this is a set of people who play cricket and suppose we know the cardinalities of these. So, let say there are 20 people who play football and 25 people who play cricket and we know that let us say 7 people who play both football and cricket.

So, we want to know how many people play at least one sport. We want to find out the number of people who play at least one sport amongst let us say football and cricket. So, we want to basically compute the cardinality of a certain set. If we denote the football players as F and the cricket players is C we want F union C and we want to count the size of that and we know that this is going to be equal to number of people who plays football plus number of people who play cricket but here we have double counted the people who play both the sport so we need to subtract that number.

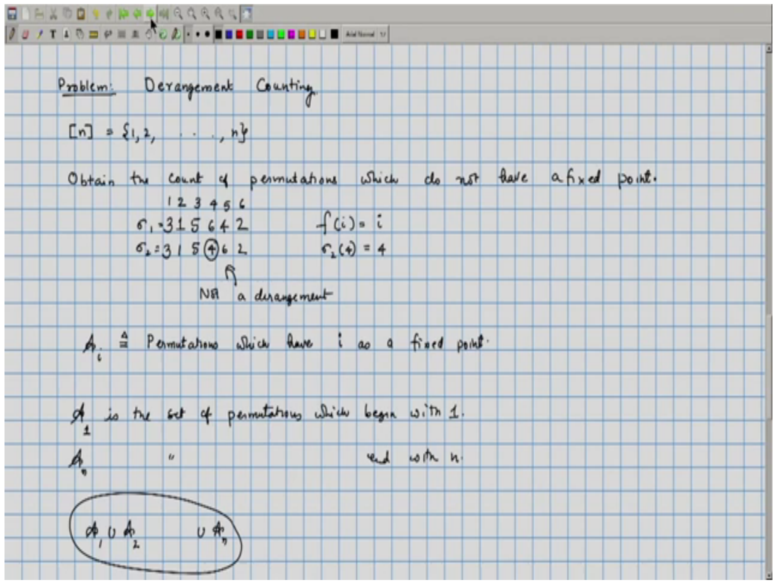
So, F intersection C whatever is the number that has to be subtracted therefore we will get 20 plus 25 minus 7 and that is going to be 38. Now, the in this case we had just two games namely football and cricket but we could have a more complex thing. Let us take a little more complex example where there are three sports you just call them as A, B and C. So, this is the set of people who play A, this is the set of people who play B, the third collection is the people who play the game C.

Now, there are intersection regions the red region is A intersection C, the orange region is A intersection B and the blue region is B intersection C and the middle region that is going to be A intersection B intersection C. As usual we want to compute the size of A union B Union C. If we just add the sizes of A, B and C; A plus B plus C the elements in the intersection regions have been counted multiple times.

For example, the red intersection region they would have been counted at least twice. The pink intersection region which is also common with the red intersection region is going to be counted three times. So, all those things have to be accounted and principle of inclusion exclusion basically it gives us a way of accounting these things in a systematic manner. So, the correct formula would be the size of A union B Union C would be size A plus size B plus size C minus A intersection B we need to remove, we also need to remove B intersection C and we also need to remove C intersection A.

Once all these have been removed when we were counting A, B and C they were counted extra and those extras have been removed but right now I mean when we just remove in this particular manner the pink region has been removed thrice so they have to be reintroduced. So, that would be size of A intersection B intersection C and now we have accounted for everything so this would be the formula. In general, then we have let us say n such sets how do we compute it is size? How do we compute the size of the union of n sets, and that is what principle of inclusion exclusion helps us do.

(Refer Slide Time: 05:36)



So, let us look at a small problem and we will illustrate the principle by means of that example. So, this is known as a derangement problem. So, we need to understand what does a derangement. So, when we say square bracket n this denotes the set of numbers from 1 to n and we are interested in counting the permutations of square bracket n which satisfies certain special properties as to obtain the count of permutations which do not have a fixed point.

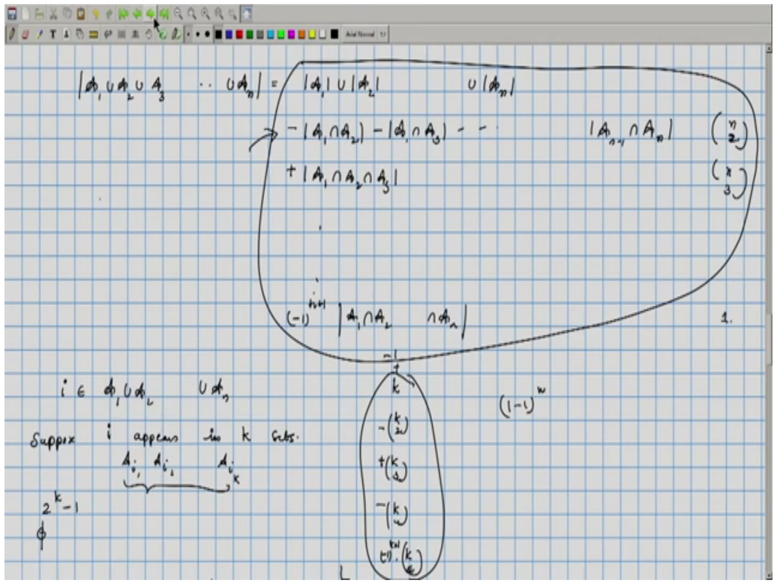
So, what is the fixed point? Let us take an example if you look at the permutation 3, 1, 5, 6, 4, 2 this is a permutation of numbers from 1 to 6, at position 1 we have 3, at position 2 we have 1, at 3 we have 5, at 4 we have 6, 5 we have 4 and at 6 we have 2. So, this could be thought of as a function from n to n any permutation would be thought of as a function from n to n such that it is not it is a bijective function.

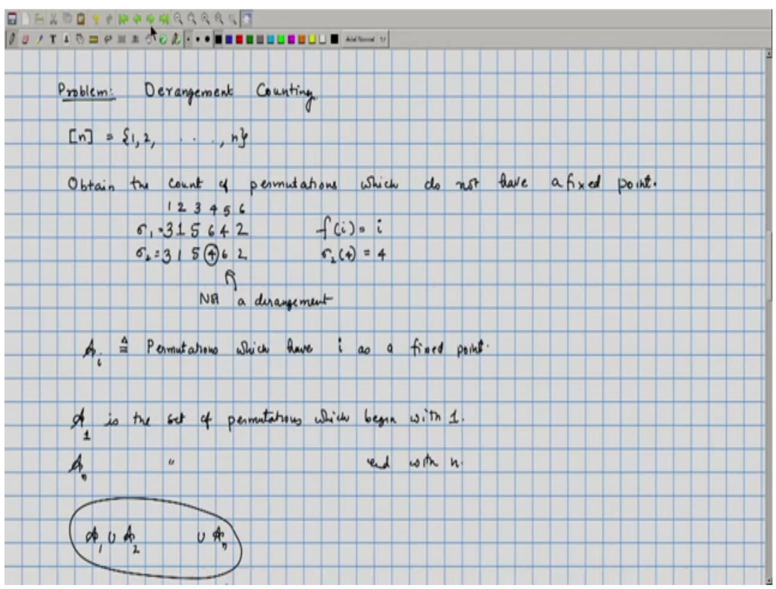
Now, in this bijective function there is no point such that f i is equal to i, whereas if we have taken the permutation 3, 1, 5, 4, 6, 2 so let us say this is sigma 1, sigma 2 is 3, 1, 5, 4, 6, 2 the 4 has a property that f sigma 2 of 4 review the permutation is a function sigma 2 of 4 is equal to 4 so this will not be a derangement. So, when we are looking at derangement, derangement is nothing but all those permutations such that at the ith position i do not appear for any value of i there are how many such permutations and that is what we need to count.

So, the approach is the total number of permutations is n factorial from this we will remove all the bad permutations instead of counting the number of derangements we will look at the complement set and we will try to count the complement and the complement sets count once obtained when that is subtracted from n factorial we will get the number of derangements. So, let us introduce some definitions.

So, let us say A i is defined as permutations there is a set of permutation which have i as a fixed point. So, fix an i so A 1 would be A 1 would basically be all the permutations with start with 1 and A n will be the set of all permutations which end with n and similarly for every other thing. And the set A 1 union A 2 union A n is a set of all permutations which are not the arrangements any derangement is basically outside this particular collection. In fact, every collection that is outside the union of A is will be at derangement. So, we need to estimate this size of A 1 union A 2 union A n. So, that is what we would do.

(Refer Slide Time: 10:10)





So, how do we compute A 1 union A 2 union A 3 all the way up to and this is where principle of inclusion exclusion comes in. So, we need to obtain the size of this. So, clearly if we just take the sizes of As and add up we are going to over count. So, A 1 union A 2 so here when we are discussing the size of the unions we need not really think of these Ais as subsets of permutations but you can think of it as the general I mean any collection of sets A 1 to A n.

So, A 1 union A 2 union A n this is going to be a an over count so we need to exclude. So, here we have included many things now we need to exclude things which has been counted twice. So, minus A 1 intersection A 2 minus A 1 intersection A 3 all the way up to A n minus 1 intersection A n note that there are going to be n choose 2 terms here look at every pair which I mean every pair of sets from A 1 to A n and take their intersection everything in the intersection would have been counted twice ones for each A i so all those things have to be removed.

But now this may be removing too many things because the (())(11:53) elements which have been which lies in the intersection of three elements so they have to be reintroduced. So, that gives us the next term A 1 intersection A 2 intersection A 3 and we need to look at and this comes with a positive sign because here in the second step we had removed too many elements so some of them we have to reintroduce and this process keeps on going.

So, n choose 3 terms would be rare and finally depending upon the sign so minus 1 raise to n plus 1 A 1 intersection A 2 intersection A n and there would be precisely one term there. Now, this is an unwieldy formula looks very complicated so we will try to write the formula in a much more nicer manner. So, what are we really doing here? And how do we know that this formula is really correct so we can look at each element?

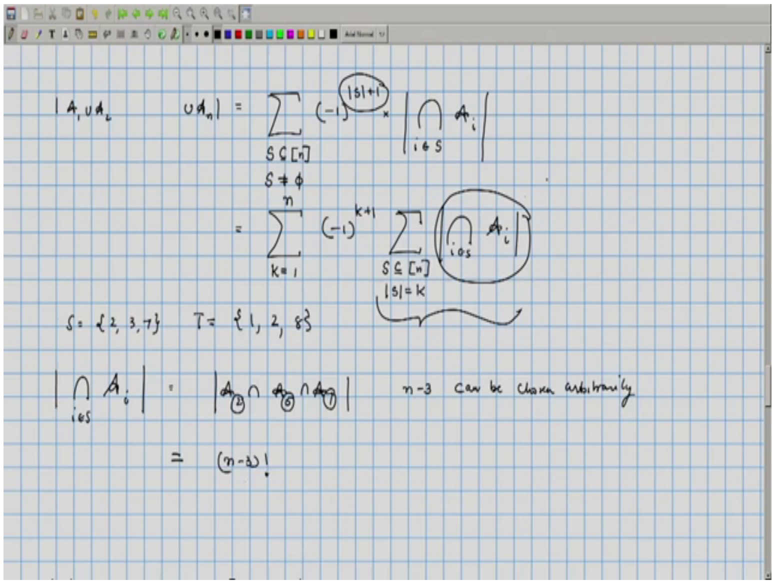
Let us say i belonging to A 1 union A 2 union A n. In the left hand side each of those elements is counted exactly once and we need to say that in the right side also each term is going to be counted exactly once. Now, which all are suppose the element i appears in k sets. Suppose i is belonging to k different sets from A 1 to A n. Only those terms are going to contribute to the count of i in the right hand side.

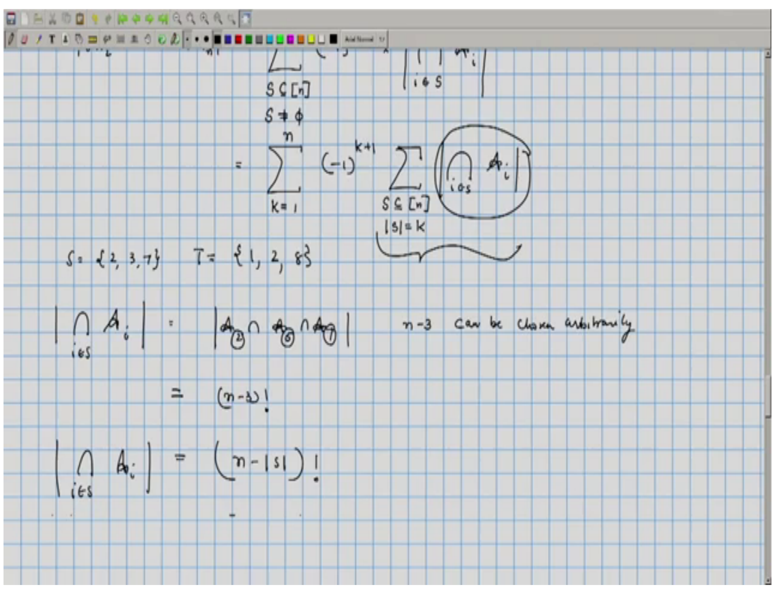
So, you can see that in the first expression the first line there would be k terms and from the second line we k choose 2 terms and so on. So, if you add this up you can see that they will add up to 1. So, the signs come appropriately so k and minus k choose 2 plus k choose 3 minus k choose 4 and so on because you take any subset of these. So, let us say i appears in A i 1, A i 2 and A i k the terms corresponding to that would be I mean in the right hand side will be all those subsets of these k sets.

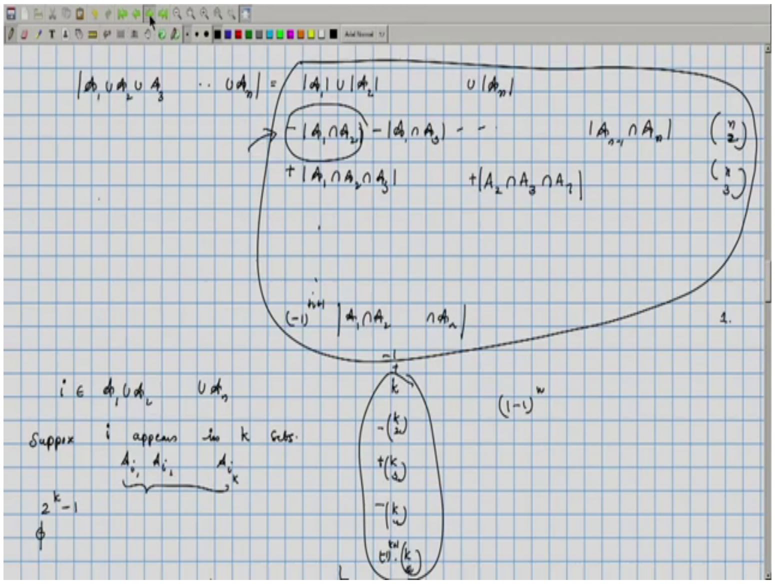
There are 2 to the k subsets and if you ignore the empty subset there are 2 to the k minus 1 subsets and their count would appear as k, k choose 1, k chose 2 all the way up to minus 1 raise to k plus 1, k choose k and if you add these up you can see that the count would be 1 because 1 plus and the sign would adjust properly if we had taken minus 1 along with it, minus 1 plus ; minus k choose 2 and so on those are the binomial coefficients when you consider the expansion of 1 minus 1 raise to n.

So, since that sum is 0 these other terms must act 1. So, that is the proof of why the inclusion exclusion principle is correct. Now, what we need to look at is how can we write this large expression in a nicer format. So, what are we really doing here? Here there are roughly 2 to the n subsets except the empty subset of 1 to n every other subset is appearing here, whereas on the left hand side there are just n terms. So, that is why the formula looks really complicated.

(Refer Slide Time: 16:20)







So, we will try to write it in a simpler way, what we will do is we will organize the subsets in terms of their size. So, size of A 1 union A 2 union A n this is basically this sum over subsets of n. We are going to associate each term on the right hand side with a subset of 1 to n and we will ignore the subset S is equal to phi. For example, the term A 1 intersection A 2 is corresponding to the subset 1 comma 2, the term A 2 intersection A 3 intersection A 7 this is going to appear somewhere with a plus sign on the third row that is going to correspond to the subset 2, 3, 7 and note that when we are considering a subset of size 2 the sign is negative and whenever we are considering a subset of size 3 it is associated sign is plus.

So, there is going to be a term minus 1 raised to size of the subset plus 1 you can also write minus 1 does not really matter only the parity of this number counts S plus 1 or S minus 1 they are I mean either both of them are even or both them are odd and this times the next term corresponds to the number of elements in the intersection. So, we will write that as so intersection over i belonging to S A i and we need to look at the size of this set.

So, the earlier expression that we had written for inclusion exclusion principle if we choose the correct notation it becomes a simplified notation. Now, this again is a sum over 2 to the n minus 1 terms we could also for some purposes enumerate the subsets in terms of their size so then you will get a double summation. So, that is we will first look at all subsets of size 1 and then look at all subsets of size 2 and so on and when we do that this term size of S plus 1 basically becomes a fixed quantity.

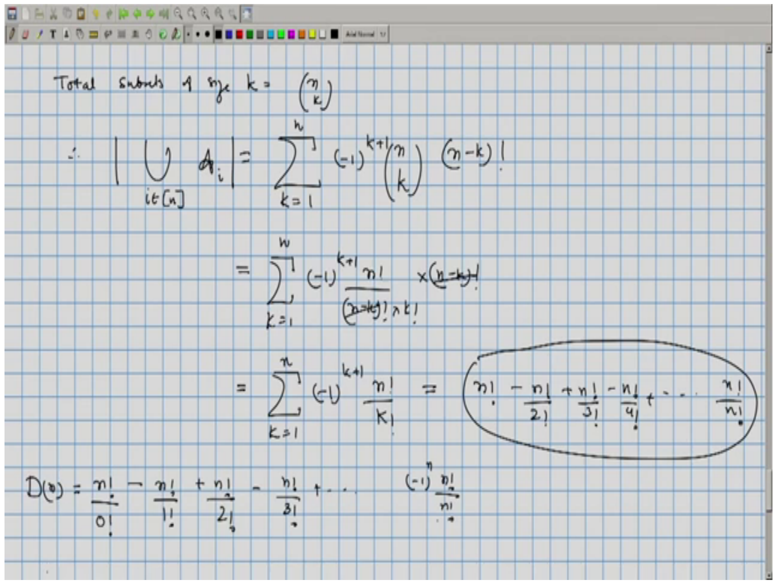
So, this would be sum over k going from 1 to n here k is going to be the size of the set and minus 1 raise to k plus 1 and then summation again now all subsets of n but now there is an additional constraint that size of S is equal to k intersection i belonging to S A i, the size of this set. So, this is a much more succinct expression than the previous expression. The meaning is the same the intuitions are also one of the same.

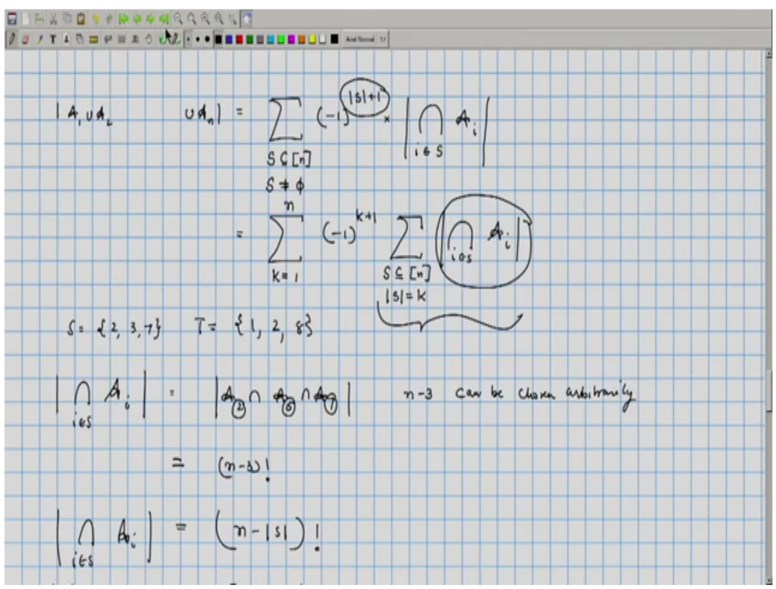
So, now so this is the general principle of inclusion and exclusion. Now, we can look at our specific case where the A is where subsets of permutation and what we need to look at is basically the inner terms of this expression. So, if we look at the innermost expression which is the intersection of the A is. So, what is the size of that going to be intersection over i belonging to S A i.

So, for example if S was the set 2, 3, 7 then this corresponds to A 2 intersection A 3 intersection A 7 and we need to look at this size of the set. Now, what really is the set? It is just all those permutations such that 2 appears at the second position, 3 appears in the third position and 7 appears at the seventh position. The other positions, if there were n numbers to choose from the other n minus 3 positions can be chosen arbitrarily.

So, this size would be n minus 3 factorial and this depends only on the size of the set it does not really depend on which are the elements that contribute to that particular set. For example, if we had instead of S we have taken another set T is equal to say 1, 2 and 8 that would also have given the exact same count. So, the inner term is going to be intersection i belonging to S A i it is size is just going to be n minus size of S factorial and the total number of such sets we need to sum up over all possible sets.

(Refer Slide Time: 21:49)





The total number of sets is going to be subsets of size k is equal to n choose k therefore the size of the union i belonging to n A i this is what we need to determine this is equal to summation minus 1 raise to k is going from 1 to n, there are n choose k subsets of size k and for each of those subsets the summation, the size of the intersection is going to be n choose it is going to be n minus k factorial.

So, this gives us summation k equals 1 to n minus 1 raise to k n factorial by n minus k factorial into k factorial the whole multiplied by n minus k factorial and that simplifies to n factorial by k factorial 1 to n minus 1 raised to n, n factorial by k factorial. So, this can simply be written as when k equals 1 this is n factorial sorry this term is k plus 1 so n factorial minus n factorial by 2 factorial plus n factorial by 3 factorial minus n factorial by 4 factorial and so on.

So, this will go on till n factorial by n factorial. So, this is the bad permutations in the sense these are the permutations that we need to avoid. So, total number of derangement would be just n factorial minus this number. So, if we denote the derangement by D n. So, D n is equal to n factorial let us just write this as n factorial by 0 factorial minus n factorial by 1 factorial plus n factorial by 2 factorial minus n factorial by 3 factorial and so on all the way up to minus 1 raised to n, n factorial by n factorial. So, that concludes the discussion on inclusion and exclusion