[SQUEAKING]

[RUSTLING]

[CLICKING]

YUFEI ZHAO: In this video, we'll look at an application of the probabilistic method to graph theory. An independent set in a graph is a subset of vertices with no two adjacent. For example, if the graph is this cycle on four vertices, an example of an independent set would be two vertices like this. And these two vertices do not-- these two vertices are not adjacent to each other, whereas had I chosen two vertices on the same edge, that would not be an independent set.

So an important question in graph theory is given a graph, what can you say about the size of its independent sets? The following theorem, due to Caro-Wei, says that every graph G contains a large independent set in the following sense. It contains an independent set of size at least the following quantity, summing over all v among vertices G, 1 over the degree of v plus 1.

So let us prove this theorem first, and then we'll see an application and some ways to interpret this result. The proof of this theorem applies the probabilistic method. So we are given this graph G. And the first thing we'll do is order the vertices of G uniformly at random.

So consider a random permutation of the vertices of G. And let's consider the set I of vertices defined as follows. These are the vertices in G such that v appears before all it's neighbors in the random ordering.

For example, from the graph earlier, had we ordered the vertices in this order-- so that's, say, some random ordering. So let's order the vertices 1, 2, 3, and 4. So the edges are 1 to 2, 2 to 4, 2 to 4, 3 to 4, and 1 to 3. What I would pick out is a set of vertices such that-- so let's see if each vertex would belong to I.

The first vertex appears before all of its neighbors. So we put the first vertex in I. The second vertex does not appear before one of its neighbors, 1, so we do not put it in I. And the third vertex here also does not appear before its neighbor. And we do not put it in I. And the fourth vertex, likewise, does not appear before all of its neighbors. And we do not put it in I as well.

So in this case, this set picks up only the first vertex. But in more complicated graphs, it can pick up additional vertices. I claim that I is always an independent set. Indeed, it will be impossible to have two vertices in I that are adjacent to each other because one of them would appear before the other in the order, and that would violate the condition on how we chose I.

Next, let's think about how large I is. For every vertex V and G, the probability that V is in I-- well, this is the probability that the V appears first among V and its neighbors.

Well v has D sub V with degree of V neighbors. And all the ordering are chosen uniformly at random, so this probability is 1 divided by the degree of V plus 1. And thus by the linearity of expectations, the expected size of this set I is equal to the sum over vertices V in G of the probability that this V lies in I. And we just computed this probability as such.

This is the expectation. This is what happens on average. And therefore, there must be some ordering that induces an I so there is some I with the size of I at least the quantity that we just produced.

So that finishes the proof of the Caro-Wei theorem. Again, it says-- the way to think about this theorem is that if you are given a graph g such that typically one does not have large degrees, then it must contain a large independent set. By considering the complement of this graph G-- by considering the graph complement-- the graph complement means that we switch, flip the edge and non-edges.

So this is a graph G. And then this will be the complement of G, So flip the edges and non-edges on G. Independent sets in G become cliques in the complement. The cliques are complete graphs with all the edges present and vice versa.

So we can read Caro-Wei's theorem for the complement of G to deduce the following corollary. It says that every n-vertex. graph G contains a clique on at least this many vertices. So the expression inside the summation is the clique of-- is the complement or the degree in the complement.

Let us derive a interesting corollary of this corollary. And this turns out to be a pretty foundational result in graph theory known as Turan's theorem. Turan's theorem says the following.

If an n-vertex graph has more than the following number, 1 minus 1 over r n squared over 2-- so it has more than this many edges-- then the graph contains clique on more than r vertices.

So Turan's theorem gives us some bound on the maximum number of edges that a graph can have if it doesn't have a large clique. And furthermore, this bound here is best possible in the following sense.

And here let us restrict to the case when n is divisible by r. When n is not divisible by r, you need to do a similar construction, turns out to be best possible, and this bound can be improved slightly, but not by too much.

So here's the example. Let us take the n vertices and split them into equal parts. And there are r parts. So here in this illustration r equals to 3. And let's put in all the edges between parts, but no edge within a part. So this is called the complete r partite graph.

So you see that this graph here, one can do a quick calculation to see that it has exactly this number, 1 minus 1 over r times n squared over 2 edges. On the other hand, it does not contain a clique on more than r vertices. The biggest clique you have comes from taking one vertex in each part. So this is an example showing that the statement of Turan's theorem is best possible.

Let us prove Turan's theorem as a quick corollary of the earlier result. So by the earlier corollary, G has a clique of size, at least sum over the vertices in G, 1 over n minus the degree of v.

And here, let us use the convexity. So in this step, let us use the convexity of the function, sending x to 1 over n minus x. So this is a convex function, which allows us to lower bound the sum by n over n minus the average degree of G.

On the other hand, if the graph has more than this many edges, then the average degree is bigger than some corresponding quantity. And putting this into the expression, we can have a final expression strictly bigger than n over n minus 1 minus 1 over r times n. And this quantity here equals to r. So there's a strict equality sign here. In particular, G has a clique of size strictly greater than r, and that finishes the proof of Turan's theorem.

Turan's theorem is a important and foundational result in graph theory. And it's a start of a subject that we now call extremal graph theory, concerning what can you say about a graph that has certain properties, like not having a large clique? And this proof that we just saw is a beautiful illustration of the probabilistic method to graph theory that allows us to prove this wonderful result, Turan's theorem, along with many other results that I hope that you will learn in the future.