Exercise session 13: The probabilistic method, edge-colourings, and the Rado graph · 1MA170

Vilhelm Agdur[1]

[1] vilhelm.agdur@math.uu.se

29 November 2023

> We introduce the probabilistic method and the Erdős-Rényi graph. Then we discuss the idea of edge-colourings of graphs, and the most basic notions of Ramsey theory. Finally, we discuss the Rado graph.

Random graphs and the probabilistic method

We have already seen one example of the probabilistic method, when we proved the result of Caro-Wei about the existence of independent sets of size at least $\sum_{v \in V}(d_v + 1)^{-1}$. We are going to see several more results from various parts of graph theory proven using the probabilistic method.

We will, in a few of the proofs, need to use a randomly chosen graph. The study of random graphs is in fact in itself a major research area in graph theory.[2] One particularly well-studied random graph is the perhaps simplest one, the *Erdős-Rényi graph*.

[2] It is also one of the main things done by graph theory people in our department, actually.

Definition 1. For any integer $n \in \mathbb{N}$ and any probability $p \in [0, 1]$, the *Erdős-Rényi graph* $G(n, p)$ is a random graph on n vertices,[3] where each of the $\binom{n}{2}$ potential edges is present independently at random with probability p.

[3] We will normally assume its vertex set is $[n]$, unless otherwise stated.

There are, in a broad overview, three kinds of argument that we tend to make when using the probabilistic method. Let us give sketches of the three methods here, and exercises using the first two methods.

Probabilistic method proof of existence: We wish to show that there exists an object with some certain property, such as a graph with high chromatic number containing no short cycles. In order to prove this, we

1. first carefully select a way to choose a random graph. If we are lucky, we can do this by choosing the right values of n and p in an Erdős-Rényi graph.

2. Then we lower bound the probability that our random graph will have the desired property – so in this case that it will have high chromatic number and no short cycles. What we need to show is that this probability is greater than zero.

3. We conclude that if the probability of an event is greater than zero, there must be some particular outcome in which the event happens

– that is, there exists such a graph.

A common trick we end up using is the so called **union bound**. We want to find an upper bound for the probability of some event A – say, the event that a random graph has a four-cycle. We divide this event up into a bunch of not necessarily mutually exclusive events $A_1, A_2, \ldots, A_n$, writing $A = \bigcup_{i=1}^n A_i$ – for example, one event for each four-tuple of vertices, where the event is that this tuple in particular forms a four-cycle. Then we use the fact that

$$\mathbb{P}\left(A\right) = \mathbb{P}\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n \mathbb{P}\left(A_i\right)$$

and bound each of the summands instead.

Exercise 1. Prove using the union bound method that if $G = G(n, p)$ is an Erdős-Rényi graph, then

$$\mathbb{P}\left(\alpha\left(G\right) \geq k\right) \leq \binom{n}{k}(1-p)^{\binom{k}{2}},$$

where $\alpha(G)$ is the independence number of G.

First moment method: We wish to show that there is an object of a certain size. For example, in Caro-Wei, we wanted to show that there is an independent set in a graph of a certain size. In order to do this, we

1. first choose a random independent set in a clever way. Again, we need to choose this in such a way that the rest of the calculations work out.

2. Then we compute the expected size of the independent set. The crucial part in this argument is normally that the expected size of a random subset A of a big set X can always be written as

$$\mathbb{E}\left[|A|\right] = \mathbb{E}\left[\sum_{x \in X} \mathbb{1}_{\{x \in A\}}\right] = \sum_{x \in X} \mathbb{E}\left[\mathbb{1}_{\{x \in A\}}\right] = \sum_{x \in X} \mathbb{P}\left(x \in A\right),$$

using linearity of expectation. So we can reduce to calculating the probability that a single vertex is in the set, which is usually much easier to do.

3. We conclude that if the expected size of the independent set is at least what we wanted, then there must be an actual outcome of this size, since the average cannot be greater than every single outcome.

Sometimes, what we want to show is not existence of something, but rather that something does *not* exist – that is, we want to show that the set of such things has size zero.[4] We can do this as well using the first moment method, replacing the last step with an application of Markov's inequality from probability theory:

[4] Normally, we mean this in an asymptotic sense – as some underlying parameter such as the number of vertices goes to infinity, the probability that the thing exists in the graph goes to zero.

Lemma 2 (Markov's inequality)**.** *Suppose X is a random variable that always takes non-negative values. Then for any $a > 0$,*

$$\mathbb{P}\left(X > a\right) \leq \frac{\mathbb{E}\left[X\right]}{a}.$$

In particular, if we have a sequence of random variables X_n such that $\mathbb{E}\left[X_n\right] \to 0$, then also $\mathbb{P}\left(X_n > 0\right) \to 0$.

Exercise 2. Recall that the *diameter* of a graph is the longest distance between any two vertices in the graph. Prove that for every $p \in (0,1)$, we have

$$\mathbb{P}\left(\mathrm{diam}\left(G(n,p)\right) = 2\right) \to 1 \qquad \text{as} \qquad n \to \infty.$$

Sometimes we want to show that something occurs basically every time in a random graph. Unfortunately, you cannot prove this just by showing that the expected number of times it occurs is large: Imagine a lottery where you have a $\frac{1}{n}$ chance to win n^2 gold bars. The expected number of gold bars you win is always n, but it is still true that the probability of winning goes to zero with n. So the expectation alone being high does not tell you you will actually usually win anything at all.

In this situation, we will instead have recourse to the second-moment method.[5]

The second moment method: We wish to show that something asymptotically almost always[6] occurs in a random graph, for example that it has a triangles. What we do is that we let X be the random number of triangles, compute the expectation and variance of X, and appeal to the following inequality[7]

$$\mathbb{P}\left(X = 0\right) \leq \frac{\mathrm{Var}\left(X\right)}{\mathbb{E}\left[X\right]^2}.$$

Edge-colourings

In our last lecture, we looked at the notion of a vertex colouring of a graph, and derived a few results about it. Of course, vertices are not the only thing we could be colouring – we could also look at colouring the *edges* of a graph.

Definition 3. Let $G = (V, E)$ be a graph. A *proper*[8] *k-edge-colouring* is a function $c : E \to [k]$ such that no two edges which are incident to each other (i.e. share an endpoint) are assigned the same colour. If we do not have this restriction on incident edges, we call it just an (improper) edge colouring.

The *edge-chromatic number* of G, denoted $\chi_1(G)$,[9] is the smallest integer k such that G has a proper k-edge-colouring.

[5] The version we give here is a slightly limited case of the more general method – in general, what makes a proof a second-moment method proof is that we compute second moments and appeal to Chebyshev's inequality.

[6] This sounds like imprecise language, but it actually is a mathematical term that means "occurs with probability converging to one".

[7] Which follows from Chebyshev's inequality, and holds whenver the right-hand side is well defined. (If the expectation is infinite, the variance is undefined, and the right hand side becomes undefined divided by infinity.)

[8] Unlike for vertex colourings, we will actually be interested in improper edge colourings more often than proper ones, so we choose the opposite convention of including the word proper and omitting the word improper for them.

[9] This is sometimes also called the *chromatic index* of G. The 1 in the notation indicates that edges are one-dimensional – if we ever need to refer to both the chromatic number and the edge-chromatic number at the same time, we may thus denote the chromatic number by $\chi_0(G)$, since vertices are zero-dimensional. In some texts the edge-chromatic number is denoted by $\chi'(G)$, but χ and χ' look way too similar in LaTeX and on a blackboard, so let us avoid that notation.

Exercise 3. Given this little discussion, can you offer a definition of what $\chi_2(G)$ might refer to for a plane graph?

Exercise 4. We observed for vertex colourings that there are trivial bounds for it in terms of the clique number and the independence number, and that each colour class is an independent set.

Now, observe that a proper edge-colouring of G is just a vertex colouring on the line graph of G. Use this observation to derive bounds on $\chi_1(G)$. What are the colour classes of a proper edge-colouring, using a term we've already defined?

In our study of vertex colourings, we proved that all planar graphs can be coloured with five colours, using a trick known as Kempe changes.[10] You can do something similar for proper edge-colourings, considering edge-induced components of edges with two colours and swapping those.

[10] We defined exactly what these are in the previous exercise session.

Exercise 5. Use a trick like this to prove the following theorem by König:[11]

[11] This exercise is probably a little bit tough, but should be doable if you give it some time. It definitely isn't as quick as the other ones could be, though.

Theorem 4 (König, 1916). *For every bipartite graph G with maximal degree Δ, we have $\chi_1(G) = \Delta$.*

Ramsey theory

Ramsey theory, named after remarkable British mathematician Frank P. Ramsey,[12] studies the question of how large a graph has to be in order to always contain a given structure. The simplest case is that of the eponymous *Ramsey number*. Let us give two definitions of it:

[12] Seriously, read his Wikipedia page if you get bored, he was almost a modern day Newton in terms of being British and inventing tons of stuff across fields.

Definition 5. For any integer k, the k-th *Ramsey number* $R(k)$ is the smallest integer n such that every edge-colouring of K_n using only two colours contains a monochromatic K_k.[13]

Definition 6. For any integer k, the k-th *Ramsey number* $R(k)$ is the smallest integer n such that every graph on n vertices contains either a k-clique or an independent set of size k.

[13] Clearly, we mean improper edge colourings here, since there is no proper 2-edge-colouring of any K_n other than K_2. By "containing a monochromatic K_k" we mean that for some colour i, the edge-induced subgraph $K_n \langle c^{-1}(i) \rangle$ contains a k-clique.

Exercise 6. Prove that the above two definitions are equivalent.

Of course, it is a non-trivial fact that these Ramsey numbers in fact even exist – a priori it could be the case that for some k, there exist arbitrarily big graphs that contain neither a k-clique nor an independent set of size k. We will prove that this is not the case in the lecture, and give a lower bound on these numbers using the probabilistic method. For now, let us just show that one Ramsey number is finite:

Exercise 7. Prove that $R(3) = 6$.

The Rado graph

In almost all of this course, we only care about finite graphs. On the few occasions we have mentioned infinite graphs, it has been to demonstrate that things can turn weird if we allow infinite graphs, because infinity is strange.

In this section, we will see perhaps the strangest example of all, the *Rado graph*, also known as THE *random graph*, with emphasis on THE.

Definition 7. The *Rado graph* is the unique[14] countably infinite homogeneous graph G such that, for any finite graph H, H is isomorphic to an induced subgraph of G.[15]

[14] Up to isomorphism.

[15] In fact, the Rado graph contains every *countably infinite* graph as an induced subgraph as well, but let's skip that in the definition.

This definition should leave you with two large questions hovering in your mind, and one smaller one:

1. Can such a thing even exist?!

2. Even if such a thing exists, how can it be unique?

3. Wait, what does it mean for a graph to be "homogeneous"?

We will explore some of the strangeness of this graph, but the definition we gave of it is a bit hard to work with. So let us offer two definitions – first, what did we mean by homogeneous?

Definition 8. Let $G = (V, E)$ be a finite or infinite graph. We say that G is *homogeneous* if, for any two subsets $A, B \subseteq V$ such that the induced subgraphs $G[A]$ and $G[B]$ are isomorphic, the isomorphism between them can be extended to an automorphism of the entire graph.

Concretely, this means that if $f : A \to B$ is the isomorphism between $G[A]$ and $G[B]$, there is an isomorphism $g : V \to V$ between G and itself, such that $f(a) = g(a)$ whenever $a \in A$.

Exercise 8. Find at least two different examples of families of homogeneous graphs.[16]

[16] There are two somewhat "trivial" examples that come to mind immediately for me, among the graph families we have already defined in the course. Plus a funny construction of an infinite graph with this property.

The version of the definition of the Rado graph that we can actually work with is in terms of *saturation*.

Definition 9. A graph $G = (V, E)$, finite or infinite, is *k-saturated* if, for any two subsets $U, W \subseteq V$, each of size at most k, there exists a vertex $v \in V$ such that $v \sim u \in E$ for every $u \in U$, and $v \sim w \notin E$ for every $w \in W$.

Exercise 9. Consider the graph G whose vertices are the two-element subsets of the set $\{1, 2, \ldots, 6\}$, and where there is an edge between two such subsets if their intersection is non-empty. Prove that this graph is 2-saturated.

Exercise 10. Prove that if G is k-saturated, then for any graph H on at most k vertices, G contains H as an induced subgraph.[17]

[17] **Hint:** Use induction in the size of H.

Having done all this, let us give the definition of the Rado graph that we will actually be using:

Definition 10. The *Rado graph* is the unique countably infinite graph which is k-saturated for every $k \in \mathbb{N}$.

Notice how the previous exercise we did proves that the Rado graph by this definition does contain every finite graph as an induced subgraph. We leave the proof of the existence of the Rado graph, and of the equivalence of the two definitions, to the actual lecture.[18]

[18] Though proving homogeneity of the Rado graph according to the second definition is actually not too hard, so if you like you can try to find a proof of this yourself.