11.3 Properties of almost all graphs

Recall that a graph property is a class of graphs that is closed under isomorphism, one that contains with every graph $G$ also the graphs isomorphic to $G$. If $p = p(n)$ is a fixed function (possibly constant), and $\mathcal{P}$ is a graph property, we may ask how the probability $P[G \in \mathcal{P}]$ behaves for $G \in \mathcal{G}(n, p)$ as $n \to \infty$. If this probability tends to 1, we say that $G \in \mathcal{P}$ for almost all (or almost every) $G \in \mathcal{G}(n, p)$, or that $G \in \mathcal{P}$ almost surely; if it tends to 0, we say that almost no $G \in \mathcal{G}(n, p)$ has the property $\mathcal{P}$.

(For example, in Lemma 11.2.1 we proved that, for a certain $p$, almost no $G \in \mathcal{G}(n, p)$ has a set of more than $\frac{1}{2}n/k$ independent vertices.)

To illustrate the new concept let us show that, for constant $p$, every fixed abstract graph $H$ is an induced subgraph of almost all graphs:

**Proposition 11.3.1.** For every constant $p \in (0, 1)$ and every graph $H$, almost every $G \in \mathcal{G}(n, p)$ contains an induced copy of $H$.

**Proof.** Let $H$ be given, and $k := |H|$. If $n \geq k$ and $U \subseteq \{0, \ldots, n-1\}$ is a fixed set of $k$ vertices of $G$, then $G[U]$ is isomorphic to $H$ with a certain probability $r > 0$. This probability $r$ depends on $p$, but not on $n$ (why not?). Now $G$ contains a collection of $\lfloor n/k \rfloor$ disjoint such sets $U$. The probability that none of the corresponding graphs $G[U]$ is isomorphic to $H$ is $(1-r)^{\lfloor n/k \rfloor}$, since these events are independent by the disjointness of the edges sets $[U]^2$. Thus

$$P[H \not\subseteq G \text{ induced}] \leq (1-r)^{\lfloor n/k \rfloor} \to 0,$$

which implies the assertion. $\square$

The following lemma is a simple device enabling us to deduce that quite a number of natural graph properties (including that of Proposition 11.3.1) are shared by almost all graphs. Given $i, j \in \mathbb{N}$, let $\mathcal{P}_{i,j}$ denote the property that the graph considered contains, for any disjoint vertex sets $U, W$ with $|U| \leq i$ and $|W| \leq j$, a vertex $v \notin U \cup W$ that is adjacent to all the vertices in $U$ but to none in $W$.

**Lemma 11.3.2.** For every constant $p \in (0, 1)$ and $i, j \in \mathbb{N}$, almost every graph $G \in \mathcal{G}(n, p)$ has the property $\mathcal{P}_{i,j}$.

---

3 The word ‘abstract’ is used to indicate that only the isomorphism type of $H$ is known or relevant, not its actual vertex and edge sets. In our context, it indicates that the word ‘subgraph’ is used in the usual sense of ‘isomorphic to a subgraph’.
11.3 Properties of almost all graphs

Proof. For fixed $U, W$ and $v \in G - (U \cup W)$, the probability that $v$ is adjacent to all the vertices in $U$ but to none in $W$, is

$$p^{|U|} q^{|W|} \geq p^i q^j.$$ 

Hence, the probability that no suitable $v$ exists for these $U$ and $W$, is

$$(1 - p^{|U|} q^{|W|})^{n - |U| - |W|} \leq (1 - p^i q^j)^{n - i - j}$$

(for $n \geq i + j$), since the corresponding events are independent for different $v$. As there are no more than $n^{i+j}$ pairs of such sets $U, W$ in $V(G)$ (encode sets $U$ of fewer than $i$ points as non-injective maps $\{0, \ldots, i - 1\} \to \{0, \ldots, n - 1\}$, etc.), the probability that some such pair has no suitable $v$ is at most

$$n^{i+j}(1 - p^i q^j)^{n - i - j},$$

which tends to zero as $n \to \infty$ since $1 - p^i q^j < 1$. □

Corollary 11.3.3. For every constant $p \in (0, 1)$ and $k \in \mathbb{N}$, almost every graph in $\mathcal{G}(n, p)$ is $k$-connected.

Proof. By Lemma 11.3.2, it is enough to show that every graph in $\mathcal{P}_{2, k-1}$ is $k$-connected. But this is easy: any graph in $\mathcal{P}_{2, k-1}$ has order at least $k + 2$, and if $W$ is a set of fewer than $k$ vertices, then by definition of $\mathcal{P}_{2, k-1}$ any other two vertices $x, y$ have a common neighbour $v \notin W$; in particular, $W$ does not separate $x$ from $y$. □

In the proof of Corollary 11.3.3, we showed substantially more than was asked for: rather than finding, for any two vertices $x, y \notin W$, some $x-y$ path avoiding $W$, we showed that $x$ and $y$ have a common neighbour outside $W$; thus, all the paths needed to establish the desired connectivity could in fact be chosen of length 2. What seemed like a clever trick in this particular proof is in fact indicative of a more fundamental phenomenon for constant edge probabilities: by an easy result in logic, any statement about graphs expressed by quantifying over vertices only (rather than over sets or sequences of vertices)\(^4\) is either almost surely true or almost surely false. All such statements, or their negations, are in fact immediate consequences of an assertion that the graph has property $\mathcal{P}_{i,j}$, for some suitable $i, j$.

As a last example of an ‘almost all’ result we now show that almost every graph has a surprisingly high chromatic number:

\[^4\]In the terminology of logic: any first order sentence in the language of graph theory
**Proposition 11.3.4.** For every constant \( p \in (0, 1) \) and every \( \epsilon > 0 \), almost every graph \( G \in \mathcal{G}(n, p) \) has chromatic number

\[
\chi(G) > \frac{\log(1/q)}{2 + \epsilon} \cdot \frac{n}{\log n}.
\]

(11.1.2) \hspace{1cm} \textbf{Proof.} For any fixed \( n \geq k \geq 2 \), Lemma 11.1.2 implies

\[
P[\alpha \geq k] \leq \binom{n}{k} q^{k(n-k)/2} \\
\leq n^k q^{k^2/2} \\
= q^{k \log n + \frac{1}{2}k(k-1)} \\
= q^{\frac{k}{2} \left( -\frac{2 \log n}{\log(1/q)} + k - 1 \right)}.
\]

For

\[
k := (2 + \epsilon) \frac{\log n}{\log(1/q)}
\]

the exponent of this expression tends to infinity with \( n \), so the expression itself tends to zero. Hence, almost every \( G \in \mathcal{G}(n, p) \) is such that in any vertex colouring of \( G \) no \( k \) vertices can have the same colour, so every colouring uses more than

\[
\frac{n}{k} = \frac{\log(1/q)}{2 + \epsilon} \cdot \frac{n}{\log n}
\]
colours. \qed

By a result of Bollobás (1988), Proposition 11.3.4 is sharp in the following sense: if we replace \( \epsilon \) by \(-\epsilon\), then the lower bound given for \( \chi \) turns into an upper bound.

Most of the results of this section have the interesting common feature that the values of \( p \) played no role whatsoever: if almost every graph in \( \mathcal{G}(n, \frac{1}{2}) \) had the property considered, then the same was true for almost every graph in \( \mathcal{G}(n, 1/1000) \). How could this happen?

Such insensitivity of our random model to changes of \( p \) was certainly not intended: after all, among all the graphs with a certain property \( \mathcal{P} \) it is often those having \( \mathcal{P} \) ‘only just’ that are the most interesting—for those graphs are most likely to have different properties too, properties to which \( \mathcal{P} \) might thus be set in relation. (The proof of Erdős’s theorem is a good example.) For most properties, however—and this explains the above phenomenon—the critical order of magnitude of \( p \) around which the property will ‘just’ occur or not occur lies far below any constant value of \( p \): it is typically a function of \( n \) tending to zero as \( n \to \infty \).
Let us then see what happens if \( p \) is allowed to vary with \( n \). Almost immediately, a fascinating picture unfolds. For edge probabilities \( p \) whose order of magnitude lies below \( n^{-2} \), a random graph \( G \in \mathcal{G}(n,p) \) almost surely has no edges at all. As \( p \) grows, \( G \) acquires more and more structure: from about \( p = \sqrt{n} n^{-2} \) onwards, it almost surely has a component with more than two vertices, these components grow into trees, and around \( p = n^{-1} \) the first cycles are born. Soon, some of these will have several crossing chords, making the graph non-planar. At the same time, one component outgrows the others, until it devours them around \( p = (\log n)n^{-1} \), making the graph connected. Hardly later, at \( p = (1 + \epsilon)(\log n)n^{-1} \), our graph almost surely has a Hamilton cycle!

It has become customary to compare this development of random graphs as \( p \) grows to the evolution of an organism: for each \( p = p(n) \), one thinks of the properties shared by almost all graphs in \( \mathcal{G}(n,p) \) as properties of ‘the’ typical random graph \( G \in \mathcal{G}(n,p) \), and studies how \( G \) changes its features with the growth rate of \( p \). As with other species, the evolution of random graphs happens in relatively sudden jumps: the critical edge probabilities mentioned above are thresholds below which almost no graph and above which almost every graph has the property considered. More precisely, we call a real function \( t = t(n) \) with \( t(n) \neq 0 \) for all \( n \) a threshold function for a graph property \( \mathcal{P} \) if the following holds for all \( p = p(n) \), and \( G \in \mathcal{G}(n,p) \):

\[
\lim_{n \to \infty} P[ G \in \mathcal{P} ] = \begin{cases} 
0 & \text{if } p/t \to 0 \text{ as } n \to \infty \\
1 & \text{if } p/t \to \infty \text{ as } n \to \infty.
\end{cases}
\]

If \( \mathcal{P} \) has a threshold function \( t \), then clearly any positive multiple \( ct \) of \( t \) is also a threshold function for \( \mathcal{P} \); thus, threshold functions in the above sense are only ever unique up to a multiplicative constant.\(^5\)

Which graph properties have threshold functions? Natural candidates for such properties are increasing ones, properties closed under the addition of edges. (Graph properties of the form \( \{ G \mid G \supseteq H \} \), with \( H \) fixed, are common increasing properties; connectedness is another.) And indeed, Bollobás & Thomason (1987) have shown that all increasing properties, trivial exceptions aside, have threshold functions.

In the next section we shall study a general method to compute threshold functions.

We finish this section with a little gem, the one and only theorem about infinite random graphs. Let \( \mathcal{G}(\aleph_0,p) \) be defined exactly like \( \mathcal{G}(n,p) \) for \( n = \aleph_0 \), as the (product) space of random graphs on \( \aleph_0 \) whose edges are chosen independently with probability \( p \).

---

\(^5\) Our notion of threshold reflects only the crudest interesting level of screening: for some properties, such as connectedness, one can define sharper thresholds where the constant factor is crucial. Note also the role of the constant factor in our comparison of connectedness with hamiltonicity in the previous paragraph.
As we saw in Lemma 11.3.2, the properties \( \mathcal{P}_{i,j} \) hold almost surely for finite random graphs with constant edge probability. It will therefore hardly come as a surprise that an infinite random graph almost surely (which now has the usual meaning of 'with probability 1') has all these properties at once. However, in Chapter 8.3 we saw that, up to isomorphism, there is exactly one countable graph, the Rado graph \( R \), that has property \( \mathcal{P}_{i,j} \) for all \( i, j \in \mathbb{N} \) simultaneously; this joint property was denoted as \((\ast)\) there. Combining these facts, we get the following rather bizarre result:

**Theorem 11.3.5.** (Erdős and Rényi 1963)
With probability 1, a random graph \( G \in \mathcal{G}(\aleph_0, p) \) with \( 0 < p < 1 \) is isomorphic to the Rado graph \( R \).

**Proof.** Given fixed disjoint finite sets \( U, W \subseteq \mathbb{N} \), the probability that a vertex \( v \notin U \cup W \) is not joined to \( U \cup W \) as expressed in property \((\ast)\) of Chapter 8.3 (i.e., is not joined to all of \( U \) or is joined to some vertex in \( W \)) is some number \( r < 1 \) depending only on \( U \) and \( W \). The probability that none of \( k \) given vertices \( v \) is joined to \( U \cup W \) as in \((\ast)\) is \( r^k \), which tends to 0 as \( k \to \infty \). Hence the probability that all the (infinitely many) vertices outside \( U \cup W \) fail to witness \((\ast)\) for these sets \( U \) and \( W \) is 0.

Now there are only countably many choices for \( U \) and \( W \) as above. Since the union of countably many sets of measure 0 again has measure 0, the probability that \((\ast)\) fails for any sets \( U \) and \( W \) is still 0. Therefore \( G \) satisfies \((\ast)\) with probability 1. By Theorem 8.3.1 this means that, almost surely, \( G \simeq R \).

How can we make sense of the paradox that the result of infinitely many independent choices can be so predictable? The answer, of course, lies in the fact that the uniqueness of \( R \) holds only up to isomorphism. Now, constructing an automorphism for an infinite graph with property \((\ast)\) is a much easier task than finding one for a finite random graph, so in this sense the uniqueness is no longer that surprising. Viewed in this way, Theorem 11.3.5 expresses not a lack of variety in infinite random graphs but rather the abundance of symmetry that glosses over this variety when the graphs \( G \in \mathcal{G}(\aleph_0, p) \) are viewed only up to isomorphism.

### 11.4 Threshold functions and second moments

Consider a graph property of the form

\[ \mathcal{P} = \{ G \mid X(G) \geq 1 \}, \]

where \( X \geq 0 \) is a random variable on \( \mathcal{G}(n, p) \). Many properties can be expressed naturally in this way; if \( X \) denotes the number of spanning trees, for example, then \( \mathcal{P} \) corresponds to connectedness.