Hereditary Properties of Graphs: Asymptotic Enumeration, Global Structure, and Colouring

Béla Bollobás

1991 Mathematics Subject Classification: Primary: 05C, 05D; secondary: 51M16, 60E15

1. Introduction. In this paper we shall discuss recent developments concerning hereditary graph properties. In particular, we shall study the growth of the number of graphs with a given hereditary property; the structure of a 'typical' graph with the property; and the $\mathcal{P}$-chromatic number of a random graph $G_{n,p}$ for a fixed hereditary property $\mathcal{P}$.

A *graph property* $\mathcal{P}$ is a union of isomorphism classes of finite graphs. To avoid trivialities, we shall always assume that our properties contain infinitely many non-isomorphic graphs, but that for some n do not contain all graphs of order n. Here are some simple examples of graph properties: (i) all triangle-free graphs without 8-cycles, (ii) all graphs of chromatic number at most k, (iii) all graphs containing no induced quadrilaterals, (iv) all regular bipartite graphs, (v) all Hamiltonian graphs.

Rather than considering general properties, we frequently study hereditary properties. A property $\mathcal{P}$ is *hereditary* if it is closed under taking induced subgraphs. In other words, $\mathcal{P}$ is hereditary if $G \in \mathcal{P}$ implies that $G - x \in \mathcal{P}$ for every vertex x of G.

An important subclass of hereditary properties is the class of *monotone properties*, those that are closed under taking subgraphs. Thus $\mathcal{P}$ is monotone if $G \in \mathcal{P}$ implies that $G - x \in \mathcal{P}$ for every vertex x of G *and* $G - e \in \mathcal{P}$ for every edge e of G. Note that properties (i) and (ii) are monotone, (iii) is hereditary but not monotone, and properties (iv) and (v) are not hereditary.

The most natural way of measuring the size of a property is to take the number of elements in its finite sections. Given a property $\mathcal{P}$, write $\mathcal{P}^n$ for the set of graphs in $\mathcal{P}$ with vertex set $[n] = \{1, \dots, n\}$. Then $(|\mathcal{P}^n|)_{n=1}^{\infty}$ is, in an obvious sense, a measure of $\mathcal{P}$.

For a monotone property $\mathcal{P}$ there is another natural measure: the sequence $(e(\mathcal{P}^n))_{n=1}^{\infty}$, where $e(\mathcal{P}^n)$ is the maximal *size* (number of edges) of a graph in $\mathcal{P}^n$. For a general property $\mathcal{P}$, the sequence $(e(\mathcal{P}^n))_{n=1}^{\infty}$ may have little significance, so we have to turn to a natural extension of it. A *pregraph* is a triple $\tilde{G} = (V, \tilde{E}, \tilde{N})$, where V is a finite set, the set of *vertices*, and $\tilde{E}$ and $\tilde{N}$ are disjoint subsets of $V^{(2)}$, the set of unordered pairs of vertices; $\tilde{E}$ is the set of *edges* and $\tilde{N}$ is the set of *non-edges* of $\tilde{G}$. A graph $G = (V, E)$ *extends* $\tilde{G}$ if $\tilde{E} \subset E \subset V^{(2)} \setminus \tilde{N}$. The *size* $e(\tilde{G})$ of a pregraph is $|V^{(2)} \setminus (\tilde{E} \cup \tilde{N})|$: the number of choices we have when extending $\tilde{G}$ to a graph. We say that a pregraph $\tilde{G}$ *belongs* to $\mathcal{P}^n$ if every graph extending $\tilde{G}$ belongs to $\mathcal{P}^n$. Then another natural measure of the size of a property $\mathcal{P}$ is the sequence $(e_n(\mathcal{P}))_{n=1}^{\infty}$, where $e_n(\mathcal{P})$ is the maximal size of a pregraph in $\mathcal{P}^n$.

It is natural to identify a graph $G = (V, E)$ with the pregraph $\tilde{G} = (V, \emptyset, V^{(2)} \setminus E)$; with this identification we find that $e(G) = e(\tilde{G})$. Hence, for a monotone property $\mathcal{P}$, the two definitions give the same value: in other words, $e(\mathcal{P}^n) = e_n(\mathcal{P})$.

Scheinerman and Zito [28] were the first to study the rate of growth of $|\mathcal{P}^n|$ for a hereditary property. They discovered that, crudely, $|\mathcal{P}^n|$ behaves in one of the following five ways: (i) for n large enough, $|\mathcal{P}^n| = 1$ or 2, (ii) it grows polynomially: for some positive integer k, $a_1 n^k \le |\mathcal{P}^n| \le a_2 n^k$ for some $a_1, a_2 > 0$, (iii) it grows exponentially: $a a_1^n \le |\mathcal{P}^n| \le a_2^n$ for some $a > 0$ and $1 < a_1 \le a_2$, (iv) it grows factorially: $a n^{a_1 n} \le |\mathcal{P}^n| \le n^{a_2 n}$ for some $a > 0$ and $0 < a_1 \le a_2$, (v) it grows superfactorially: $|\mathcal{P}^n| > n^{an}$ for every $a > 0$ and n large enough.

Here we are interested in properties whose rate of growth is not far from maximal. To measure the rate of growth of such a property $\mathcal{P}$, we replace the sequence $(|\mathcal{P}^n|)_{n=1}^{\infty}$ by the sequence $(c_n)_{n=1}^{\infty}$, where $|\mathcal{P}^n| = 2^{c_n \binom{n}{2}}$. Since $1 \le |\mathcal{P}^n| \le 2^{\binom{n}{2}}$, we have $0 \le c_n \le 1$.

We call c_n the *logarithmic density* of $\mathcal{P}^n$, and $c = \lim_{n\to\infty} c_n$ the *asymptotic logarithmic density* of $\mathcal{P}$ provided the limit exists.

Similarly, the (*normalized*) *size* of $\mathcal{P}^n$ is d_n, $0 \le d_n \le 1$, defined by $e_n(\mathcal{P}) = d_n \binom{n}{2}$.

The *asymptotic size* of $\mathcal{P}$ is $d = \lim_{n\to\infty} d_n$, provided this limit exists. Since every pregraph $\tilde{G}$ extends to $2^{e(\tilde{G})}$ graphs, we have $c_n \ge d_n$ for every property. Hence if $\mathcal{P}$ is a property with asymptotic logarithmic density c and asymptotic size d, then $c \ge d$. We shall see later that every hereditary property has an asymptotic logarithmic density c and an asymptotic size d and, in fact, they are equal.

2. Monotone properties. One of the main aims of classical extremal graph theory is the study of the sequence $(e_n(\mathcal{P}))_{n=1}^{\infty}$ for various monotone graph properties. Frequently, a monotone property is given by a family $\mathcal{F}$ of forbidden subgraphs. For a family $\mathcal{F} = \{F_1, F_2, \dots\}$ of finite graphs, let $\mathrm{Mon}(\mathcal{F})$ be the collection of all graphs containing no F_i as a subgraph. Clearly every monotone property is of the form $\mathrm{Mon}(\mathcal{F})$ for some family $\mathcal{F}$ of *forbidden subgraphs*, but one is especially interested in monotone families defined by small families of forbidden subgraphs. If there is only one forbidden subgraph F then we have a *principal* monotone property and we write $\mathrm{Mon}(F)$ instead of $\mathrm{Mon}(\{F\})$.

It has been known for over fifty years that every monotone graph property has an asymptotic size. In particular, a weak form of Turán's theorem [31] states that $d(\mathrm{Mon}(K_{r+1})) = 1 - \frac{1}{r}$ for every $r \ge 1$. The fundamental theorem of Erdős and Stone [15] extends this result to $d(\mathrm{Mon}(K_{r+1}(t))) = 1 - \frac{1}{r}$ for all $r, t \ge 1$. Here, as usual, K_n denotes a complete graph of order n and $K_r(t)$ denotes the complete r-partite graph in which each part has t vertices. An equivalent form of the Erdős-Stone theorem is that if $\mathcal{F}$ is any family of forbidden subgraphs then $d(\mathrm{Mon}(\mathcal{F})) = 1 - \frac{1}{r}$, where $r = \min\{\chi(F) - 1 : F \in \mathcal{F}\}$ and $\chi(F)$ is the chromatic number of F.

Rather more effort is needed to prove that every monotone property has an asymptotic density. Using the method of Kleitman and Rothschild [18], Erdős, Kleitman and Rothschild [12] proved that $c(\mathrm{Mon}(K_{r+1})) = 1 - \frac{1}{r}$. This result was extended by Erdős, Frankl and Rödl [11], who proved that $c(\mathrm{Mon}(F)) = 1 - \frac{1}{r}$ for every graph F, where $r = \chi(F) - 1$. The proof of this result implies that $c(\mathrm{Mon}(\mathcal{F})) = 1 - \frac{1}{r}$ for every family $\mathcal{F}$, where r is, as before, one smaller than the minimal chromatic number of a graph in $\mathcal{F}$. In particular, $c(\mathcal{P}) = d(\mathcal{P})$ for every monotone family.

The structure of K_{r+1}-free graphs was investigated in great detail by Kolaitis, Prőmel and Rothschild [19]. Among other results, they proved that $\mathrm{Mon}(K_{r+1})$ is well approximated by the smaller property $\mathcal{N}_r$ of graphs of chromatic number at most r: not only do we have the crude result that $c(\mathrm{Mon}(K_{r+1})) = c(\mathcal{N}_r)$, but also

$$|\mathrm{Mon}(K_{r+1})^n|/|\mathcal{N}_r^n| = 1 + O(n^{-k})$$

for all $k > 0$. Furthermore, a first-order labelled $0 - 1$ law holds for the class of K_{r+1}-free graphs.

Before leaving monotone properties, let us note that the following somewhat surprising fact is an immediate consequence of the description of $c(\mathcal{P}) = d(\mathcal{P})$ for a monotone property. If $\mathcal{P}_1$ and $\mathcal{P}_2$ are monotone properties, and $\mathcal{P} = \mathcal{P}_1 \cap \mathcal{P}_2$, then

$$c(\mathcal{P}) = \min\{c(\mathcal{P}_1), c(\mathcal{P}_2)\}. \tag{1}$$

Thus the intersection of two monotone properties is about as large as the smaller of the two properties!

3. VOLUMES OF PROJECTIONS AND ASYMPTOTIC ENUMERATION. The existence of the asymptotic logarithmic density of a hereditary property is closely related to a family of inequalities involving volumes of projections of bodies. Our next aim is to describe this relationship.

A *body* in $\mathbf{R}^n$ is a compact convex subset of $\mathbf{R}^n$ that is the closure of its interior. Let $v_1, \dots, v_n$ be the standard basis of $\mathbf{R}^n = \mathrm{lin}\{v_1, \dots, v_n\}$. For a subset A of $[n]$, write K_A for the orthogonal projection of a body K onto $\mathrm{lin}\{v_j : j \in A\}$, and $|K_A|$ for the $|A|$-*dimensional volume* of K_A. In particular, $|K| = |K_{[n]}|$ is the volume of K. With $\beta(K) = (|K_A| : A \subset [n]) = (|K_A|)_{A \subset [n]} \in \mathbf{R}^{\mathcal{P}(n)} = \mathbf{R}^{2^n}$, the map $K \to \beta(K)$ can be considered to be a measure of the size of the boundary of K.

We are interested in the best possible isoperimetric inequalities involving the boundary vector $\beta(K)$ and the volume $|K|$. In other words, we would like to know for which vectors $(x_A) \in \mathbf{R}^{2^n}$ with $x_{[n]} = 1$ is there a body $K \subset \mathbf{R}^n$ of volume 1 such that $|K_A| \le x_A$ for all $A \subset [n]$. The following *box theorem* we proved with Thomason [6] gives a surprisingly simple answer to this question. A *box* B in $\mathbf{R}^n$ is a body of the form $B = \prod_{j=1}^n I_j$, where each I_j is an interval.

THEOREM 1. *For every body $K \subset \mathbf{R}^n$, there is a box $B \subset \mathbf{R}^n$ such that $|B| = |K|$ and $|B_A| \le |K_A|$ for every $A \subset [n]$.*

An immediate consequence of the box theorem is the uniform cover inequality below, extending the Loomis-Whitney inequality [20]. A sequence $(A_i)_{i=1}^m$ of subsets of $[n]$ is a *k-uniform cover* of $[n]$ if every element of $[n]$ belongs to precisely k of the sets $A_1, \ldots, A_m$. Now, if $(A_i)_{i=1}^m$ is a k-uniform cover of $[n]$, and K is a body in $\mathbf{R}^n$ then Theorem 1 implies that

$$|K|^k \le \prod_{i=1}^{m} |K_{A_i}|. \tag{2}$$

In fact, in [6] the box theorem is deduced from the uniform cover inequality (1) by a simple compactness argument. Since the original proof, several other deductions have been suggested: Ball noted that separation theorems, and Kahn and Meshulam pointed out that properties of submodular functions, can be used to deduce the box theorem from inequality (2).

The box theorem easily implies that, as first proved by Alekseev [1], every hereditary property of graphs has an asymptotic logarithmic density.

THEOREM 2. *Let $\mathcal{P}$ be a hereditary property of graphs. Then $1 = c_1(\mathcal{P}) \ge c_2(\mathcal{P}) \ge \cdots$; in particular, the asymptotic logarithmic density $c(\mathcal{P}) = \lim_{n\to\infty} c_n(\mathcal{P})$ exists.*

It is easily seen that the arguments above apply to hereditary properties of r-uniform hypergraphs as well, *mutatis mutandis*.

4. ASYMPTOTIC ENUMERATION AND GLOBAL STRUCTURE. Given a family $\mathcal{F} = \{F_1, F_2, \ldots\}$, of finite graphs, let $\mathrm{Her}(\mathcal{F})$ be the collection of all graphs that contain no F_i as an *induced* subgraph. Clearly, every hereditary property is of the form $\mathrm{Her}(\mathcal{F})$ for some family $\mathcal{F}$ of *forbidden subgraphs*. Theorem 2 tells us that every hereditary property $\mathcal{P} = \mathrm{Her}(\mathcal{F})$ has an asymptotic logarithmic density $c(\mathcal{P})$, but gives no indication as to how one could determine $c(\mathcal{P})$ from $\mathcal{F}$. In fact, Prömel and Steger [22], [23], [24], [25] gave such a description for a principal hereditary property, i.e., for one with a single forbidden induced subgraph. They also gave approximations of principal hereditary properties by rather simple (non-principal) hereditary properties. With Thomason [7] we extended these results to general hereditary properties.

Before we can describe these results, we have to introduce some definitions. An (r,s)-*colouring* of a graph $G = (V,E)$ is a partition of the vertex set into r classes such that the first s classes induce complete graphs, and the remaining $r-s$ classes induce empty subgraphs. (Needless to say, empty classes are allowed.) Thus an $(r,0)$-colouring of a graph is precisely a standard r-colouring. We write $\mathcal{P}_{r,s}$ for the collection of all (r,s)-colourable graphs; clearly, $\mathcal{P}_{r,s}$ is a hereditary property for all $0 \le s \le r$, $r \ge 1$. For example, $\mathcal{P}_{1,1}$ is the collection of all complete graphs, and $\mathcal{P}_{1,0}$ is the collection of all empty graphs. The *colouring number* $r(\mathcal{P})$ of a property $\mathcal{P}$ is

$$r(\mathcal{P}) = \max\{r : \mathcal{P}_{r,s} \subset \mathcal{P} \text{ for some } s\}.$$

Note that if $\mathcal{P} = \mathrm{Her}(\mathcal{F})$ then

$$r(\mathcal{P}) = \max\{r : \text{ for some } s \le r, \text{ no } F \in \mathcal{F} \text{ is } (r,s)\text{-colourable}\}.$$

If $\mathcal{P} = \mathrm{Mon}(\mathcal{F})$ then $r(\mathcal{P})$ is exactly as before:

$$r(\mathcal{P}) = \min\{\chi(F) - 1 : F \in \mathcal{F}\} = \max\{r : \text{ no } F \in \mathcal{F} \text{ is } (r,0)\text{-colourable}\}.$$

The colouring number gives us a lower bound for c_n and d_n. Indeed, let $0 \leq s \leq r$ be such that $r = r(\mathcal{P})$ and $\mathcal{P}_{r,s} \subset \mathcal{P}$, and let $\tilde{G} = ([n], \tilde{E}, \tilde{N})$ be the pregraph obtained as follows. Partition $[n]$ into r classes as equal as possible in size, $[n] = V_1 \cup \ldots \cup V_r$, say, and let $\tilde{E}$ consist of all edges within a class V_i for $0 \leq i \leq s$. Since $\mathcal{P}_{r,s} \subset \mathcal{P}$, every extension of $\tilde{G}$ belongs to $\mathcal{P}$. Consequently,

$$c_n(\mathcal{P}) \geq d_n(\mathcal{P}) \geq e(\tilde{G}) / \binom{n}{2} \geq 1 - \frac{1}{r}.$$

As shown in [7], $c(\mathcal{P})$ and $d(\mathcal{P})$ exist for every hereditary property, and these inequalities are essentially best possible.

Theorem 3. *If $\mathcal{P}$ is any hereditary property then*

$$c(\mathcal{P}) = d(\mathcal{P}) = 1 - \frac{1}{r(\mathcal{P})},$$

where $r(\mathcal{P})$ is the colouring number of $\mathcal{P}$.

The proof of this theorem is based on the three pillars of extremal graph theory: the theorems of Ramsey [26], Erdős and Stone [15], and Szemerédi [30]. One needs only the very simple case of Ramsey's theorem that the diagonal graph Ramsey function is finite: $R(k) < \infty$ for every k. On the other hand, one needs a slight extension of the Erdős-Stone theorem: for all $r, t \geq 1$ and $\epsilon > 0$ there are $\delta > 0$ and $n_0 \in \mathbf{N}$ such that if F and G are graphs with $V(F) = V(G) = [n]$, $n \geq n_0$, $e(F) \leq \delta n^2$ and

$$e(G) \geq (1 - \frac{1}{r} + \epsilon)\binom{n}{2},$$

then G contains an F-avoiding $K_{r+1}(t)$. Here we say that a graph H *avoids* F if no edge of F joins two vertices of H.

The most important ingredient of the proof of Theorem 3 is *Szemerédi's uniformity lemma* [30]. Given a graph $G = (V, E)$, and subsets $A, B, \subset V$, the *density* $d(A, B)$ is defined as

$$d(A, B) = \frac{e(A, B)}{|A||B|},$$

where $e(A, B)$ is the number of A-B edges. A pair (A, B) is (ϵ, δ)-uniform if

$$|d(A', B') - d(A, B)| \leq \epsilon$$

whenever $A' \subset A$, $B' \subset B$, $|A'| \geq \delta|A|$ and $|B'| \geq \delta|B|$.

Szemerédi's uniformity lemma states that for all $\epsilon, \delta, \eta > 0$ there is an $M = M(\epsilon, \delta, \eta)$ such that the vertex set of every graph G can be partitioned into at most M sets $U_1, \ldots, U_m$ of sizes differing by at most 1, such that at least $(1 - \eta)m^2$ of the (ordered) pairs (U_i, U_j) are (ϵ, δ)-uniform.

The fewer sets $U_1, U_2, \dots$ we can take the more powerful the result is; unfortunately when $\epsilon = \delta = \eta$, all we know about $M(\epsilon,\epsilon,\epsilon)$ is that it is at most a tower of 2s of height proportional to ϵ^{-5}. As the proof of this bound seemed rather 'wasteful', for many years there had been some hope that this enormous bound could be reduced greatly. It was a great surprise when recently Gowers [17] proved the difficult result that $K(\epsilon,\delta,\eta)$ can not be less than of tower type in $1/\delta$, even when ϵ and η are kept large.

Szemerédi's uniformity lemma implies that every graph satisfying certain global conditions contains appropriate *induced* subgraphs; this is precisely how the lemma was used in the proof of Theorem 3.

The descriptions of the asymptotic logarithmic density and asymptotic size of a hereditary property provided by Theorem 3 imply that hereditary properties are much more complex than monotone ones. In particular, the simple relationship (1) fails for hereditary properties. For example, if $\mathcal{P}_1 = \mathrm{Her}(K_4)$, $\mathcal{P}_2 = \mathrm{Her}(C_7)$ and $\mathcal{P} = \mathcal{P}_1 \cap \mathcal{P}_2 = \mathrm{Her}\{K_4, C_7\}$, then $r(\mathcal{P}_1) = r(\mathcal{P}_2) = 3$ but $r(\mathcal{P}) = 2$: the intersection of two hereditary properties can be *much* smaller than either of them.

In fact, the intersection of two large hereditary properties need not even be a property in our sense: it may contain only finitely many non-isomorphic graphs. For example, if $r \geq 1$ then each of $\mathcal{P}_{r,0}$ and $\mathcal{P}_{r,r}$ has colouring number r, so that $c(\mathcal{P}_{r,0}) = c(\mathcal{P}_{r,r}) = 1 - \frac{1}{r}$, but $\mathcal{P}_{r,0} \cap \mathcal{P}_{r,r}$ consists of graphs G with $\chi(G) \leq r$ and $\chi(\bar{G}) \leq r$. In particular, $|G| \leq r^2$ for every $G \in \mathcal{P}_{r,0} \cap \mathcal{P}_{r,r}$, so $\mathcal{P}_{r,0} \cap \mathcal{P}_{r,r}$ indeed consists only of finitely many non-isomorphic graphs.

5. COLOURING RANDOM GRAPHS $G_{n,1/2}$ *with hereditary properties*. The random graph $G_{n,p}$ is a graph with vertex set $[n]$, whose edges are selected independently, with probability p. The probability space of these graphs is $\mathcal{G}(n,p)$. In particular, $\mathcal{G}(n,1/2)$ is the space of all $2^{\binom{n}{2}}$ graphs on $[n]$ with the uniform distribution.

One of the main questions left open by Erdős and Rényi when, almost forty years ago, they founded the theory of random graphs ([13], [14]; see also [5]) was the behaviour of the chromatic number of a random graph. Over 25 years later, first Shamir and Spencer [29] proved that the chromatic number of $G_{n,p}$ is highly concentrated, and then it was shown [3] that if $0 < p < 1$ is fixed and $q = 1 - p$ then

$$\chi(G_{n,p}) = (1 + o(1)) \frac{n}{2 \log_{1/q} n} \tag{3}$$

for almost every $G_{n,p}$. Substantial extensions of this result were proved by Łuczak [21], Frieze and Łuczak [16], and Alon and Krivelevich [2]. All these results use various martingale inequalities (see [4]).

For a property $\mathcal{P}$, a *$\mathcal{P}$-colouring* of a graph $G = (V,E)$ is a partition $V = V_1 \cup \ldots \cup V_k$ of the vertex set such that every class V_i induces a $\mathcal{P}$-graph: $G[V_i] \in \mathcal{P}, i = 1, \dots, k$. The *$\mathcal{P}$-chromatic number* $\chi_{\mathcal{P}}(G)$ of a graph G is the minimal number of classes in a $\mathcal{P}$-colouring of G. Thus $\chi_{\mathcal{P}_{1,0}}(G) = \chi(G)$ and $\chi_{\mathcal{P}_{1,1}}(G) = \chi(\overline{G})$. Scheinerman [27] was the first to study the $\mathcal{P}$-chromatic number of random graphs. He noted that if $\mathcal{P}$ is a hereditary property then either $\mathcal{P}_{1,0} \subset \mathcal{P}$ or $\mathcal{P}_{1,1} \subset \mathcal{P}$ so $\chi_{\mathcal{P}}(G) \leq \max\{\chi(G), \chi(\overline{G})\}$. From this it follows that $\chi_{\mathcal{P}}(G_{n,p}) = O(n \log n)$ for every fixed $0 < p < 1$ and hereditary property $\mathcal{P}$, and it is easily seen that, in fact, $\chi_{\mathcal{P}}(G_{n,p}) = \Theta(n \log n)$.

With Thomason [8] we proved an analogue of (3) for a general hereditary property, but only in the case $p = \frac{1}{2}$.

Theorem 4. *Let $\mathcal{P}$ be a non-trivial hereditary property of graphs, with colouring number $r = r(\mathcal{P})$. Then*

$$\chi_{\mathcal{P}}(G_{n,1/2}) = (\frac{1}{2r} + o(1))\frac{n}{\log_2 n}$$

for almost every $G_{n,1/2}$.

In fact, this result follows rather easily from (3) and from the facts that $c(\mathcal{P}) = 1 - \frac{1}{r}$ and that $\mathcal{P}_{r,s} \subset \mathcal{P}$ for some $s, 0 \leq s \leq r$. More precisely, $c(\mathcal{P}) = 1 - \frac{1}{r}$ implies that $\chi_{\mathcal{P}}(G_{n,1/2})$ is unlikely to be much smaller than $n/(2r \log_2 n)$, and $\mathcal{P}_{r,s} \subset \mathcal{P}$ implies that $\chi_{\mathcal{P}}(G_{n,1/2})$ is unlikely to be much larger than $n/(2r \log_2 n)$.

6. Colouring random graphs $G_{n,p}$ with hereditary properties. The accepted wisdom in the theory of random graphs is that whatever can be proved for the space $\mathcal{G}(n,p)$ with $p = 1/2$ can be proved for $\mathcal{G}(n,p)$ with any fixed p, $0 < p < 1$. This conventional wisdom is contradicted by the problem of determining $\chi_{\mathcal{P}}(G_{n,p})$! As we saw in Theorem 4, it is easy to determine $\chi_{\mathcal{P}}(G_{n,p})$ in the uniform case $p = 1/2$. However, for $p \neq 1/2$ not only does the proof collapse, but we are faced with a genuinely more complicated phenomenon, so that much more effort is needed to overcome the difficulties.

A lower bound for $\chi_{\mathcal{P}}(G_{n,p})$ is easily obtained from the following result, which is a consequence of the box theorem.

Theorem 5. *Let $\mathcal{P}$ be a hereditary graph property, let $0 < p < 1$ and let the constants $e_{k,p}(\mathcal{P})$ be defined by $\mathbf{P}(G_{k,p} \in \mathcal{P}) = 2^{-e_{k,p}(\mathcal{P})\binom{k}{2}}$. Then $e_{k,p}(\mathcal{P})$ increases with k. In particular, $e_{k,p}(\mathcal{P})$ tends to a limit $e_p(\mathcal{P})$ as $k \to \infty$. Furthermore, $e_p(\mathcal{P}) > 0$ if $\mathcal{P}$ is non-trivial, i.e., if not every graph has $\mathcal{P}$.*

Theorem 5 implies that, for $\epsilon > 0$, the expected number of induced subgraphs of order k in a random graph $G_{n,p}$ having property $\mathcal{P}$ is $o(1)$ for $k \geq (2/e_p + \epsilon)\log_2 n$, and tends to infinity for $k \leq (2/e_p - \epsilon)\log_2 n$. Consequently,

$$\chi_{\mathcal{P}}(G_{n,p}) \geq (e_p + o(1))n/(2\log_2 n) \tag{4}$$

almost surely.

It was conjectured in [8] that (4) is in fact an equality, as claimed by Theorem 4 for $p = 1/2$. Now, the proof of Theorem 4 is based on the fact that for $p = 1/2$ the constant $e_p(\mathcal{P})$ has a simple interpretation in terms of the values (r,s) for which $\mathcal{P}_{r,s} \subset \mathcal{P}$. However, for $p \neq 1/2$ this is no longer true: $e_p(\mathcal{P})$ cannot be characterized solely in terms of these values (r,s). For example, let $\mathcal{P}' = \mathcal{P}_{2,0}$ be the property of being bipartite, and let $\mathcal{P}''$ be the property of being 3-colourable, with two of the colour classes spanning complete bipartite graphs. Then $\mathcal{P}'$ and $\mathcal{P}''$ contain $\mathcal{P}_{1,0}$ and $\mathcal{P}_{2,0}$, and no other $\mathcal{P}_{r,s}$. Nevertheless, $e_p(\mathcal{P}') \neq e_p(\mathcal{P}'')$ for $p > 1/2$.

In spite of these difficulties, with Thomason [9] we proved the conjecture above.

THEOREM 6. *Let $\mathcal{P}$ be a hereditary graph property and let $0 < p < 1$. Let $e_p = e_p(\mathcal{P})$ be the constant defined in Theorem 5. Then*

$$\chi_{\mathcal{P}}(G_{n,p}) = (e_p + o(1))n/(2\log_2 n)$$

almost surely.

The proof of Theorem 6 makes use of Szemerédi's uniformity lemma, martingale inequalities and, above all, a careful study of the structure of a general hereditary property. The *product* $\prod_{\gamma\in\Gamma}\mathcal{P}_\gamma$ of hereditary properties $\mathcal{P}_\gamma$, $\gamma \in \Gamma$, is the class of graphs G with vertex sets $\bigcup_{\gamma\in\Gamma} V_\gamma$ such that $G[V_\gamma] \in \mathcal{P}_\gamma$ for every $\gamma \in \Gamma$. A hereditary property is *irreducible* if it is not the product of two other hereditary properties. It is easily shown that every hereditary property is the product of a finite collection of irreducible hereditary properties. Also, if $\mathcal{P} = \prod_{\gamma\in\Gamma}\mathcal{P}_\gamma$ then

$$e_p(\mathcal{P})^{-1} = \sum_{\gamma\in\Gamma} e_p(\mathcal{P}_\gamma)^{-1}.$$

Next, one can show that if Theorem 6 holds for each of the properties $\mathcal{P}_1, \dots, \mathcal{P}_k$, then it holds for $\prod_{i=1}^k \mathcal{P}_i$ as well. Consequently, it suffices to prove Theorem 6 for irreducible properties.

In fact, the heart of the proof is the assertion that Theorem 6 holds for every 'typed' property $\mathcal{P} = \mathcal{P}(\tau)$. A *type* is a labelled graph, each of whose vertices and edges is coloured black or white. Given a type τ, the property $\mathcal{P}(\tau)$ consists of those graphs G for which $V(G)$ has a partition $\bigcup_{t\in V(\tau)} V_t$ such that $G[V_t]$ is complete or empty according as t is black or white, and moreover, if the edge tu is in τ then $G[V_t, V_u]$ is a complete or empty bipartite graph according as the edge tu is black or white. The proof of the fact that Theorem 6 holds for typed properties $\mathcal{P}(\tau)$ is based on a careful analysis of the maximal number of induced edge-disjoint subgraphs of a given order having property $\mathcal{P}$ – after much work enough can be deduced so that martingale inequalities can be applied.

7. *Open problems.* Numerous open problems remain. Concerning graphs, all the discussion above is about rather 'rich' properties $\mathcal{P}$, namely those with $c(\mathcal{P}) > 0$. The case $c(\mathcal{P}) = 0$ is not understood nearly as well.

Although we know that $c(\mathcal{P}) = d(\mathcal{P})$ for every hereditary property, this is far from being the entire story. We always have

$$|\mathcal{P}^n| = 2^{c_n\binom{n}{2}} \geq 2^{e_n(\mathcal{P})} = 2^{d_n\binom{n}{2}},$$

but it would be good to decide whether $c_n = (1 + o(1))d_n$ holds as well.

More importantly, we know very little about hypergraphs. The quantities $c_n(\mathcal{P})$ and $d_n(\mathcal{P})$ are easily defined for r-graphs, and $c_n(\mathcal{P}) \geq d_n(\mathcal{P})$ for every n. Also, the box theorem implies that $c_n(\mathcal{P}) \to c(\mathcal{P})$, and one can show that $d_n(\mathcal{P}) \to d(\mathcal{P})$, but we do not know whether we always have $c(\mathcal{P}) = d(\mathcal{P})$. Nothing of importance is know about the $\mathcal{P}$-chromatic number of r-graphs: we do not even know the asymptotic $\mathcal{P}$-chromatic number of random r-graphs $G_{n,p}^{(r)}$ for $p = 1/2$.

References

[1] V.E. Alekseev, On the entropy values of hereditary classes of graphs, *Discrete Math. Appl.* 3 (1993), 191–199.

[2] N. Alon and M. Krivelevich, The concentration of the chromatic number of random graphs, *Combinatorica* 17 (1997), 303–313.

[3] B. Bollobás, The chromatic number of random graphs, *Combinatorica* 8 (1988), 49–55.

[4] B. Bollobás, Martingales, isoperimetric inequalities and random graphs, *in "Combinatorics (Eger, 1987)"* (A. Hajnal, L. Lovász and V.T. Sós, eds.), Colloq. Math. Soc. János Bolyai 52, North-Holland, Amsterdam (1988), 113–139.

[5] B. Bollobás, *Random Graphs*, Academic Press, London, 1985, xvi+447pp.

[6] B. Bollobás and A. Thomason, Projections of bodies and hereditary properties of hypergraphs, *J. London Math. Soc.* 27 (1995), 417–424.

[7] B. Bollobás and A. Thomason, Hereditary and monotone properties of graphs, in "The Mathematics of Paul Erdős II" (R.L. Graham and J. Nešetřil, eds.) *Algorithms and Combinatorics*, Vol. 14, Springer-Verlag (1997), 70–78.

[8] B. Bollobás and A. Thomason, Generalized chromatic numbers of random graphs, *Random Structures and Algorithms* 6 (1995), 353–356.

[9] B. Bollobás and A. Thomason, Colouring random graphs by hereditary properties, to appear.

[10] B. Bollobás and P. Erdős, On the structure of edge graphs, *Bull. London Math. Soc.* 5 (1973), 317–321.

[11] P. Erdős, P. Frankl and V. Rödl, The asymptotic enumeration of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent, *Graphs and Combinatorics* 2 (1986), 113–121.

[12] P. Erdős, D.J. Kleitman and B.L. Rothschild, Asymptotic enumeration of K_n-free graphs, in *International Coll. Comb.*, Atti dei Convegni Lincei (Rome) 17 (1976), 3–17.

[13] P. Erdős and A. Rényi, On random graphs I, *Publ. Math. Debrecen* 6 (1959), 290–297.

[14] P. Erdős and A. Rényi, On the evolution of random graphs, *Publ. Math. Inst. Hungar. Acad. Sci.* 5 (1961), 17–61.

[15] P. Erdős and A.H. Stone, On the structure of linear graphs, *Bull. Amer. Math. Soc.* 52 (1946), 1087–1091.

[16] A.M. Frieze and T. Łuczak, On the independence and chromatic numbers of random regular graphs, *J. Combinat. Theory (B)* 54 (1992), 123–132.

[17] W.T. Gowers, Lower bounds of tower type for Szemerédi's uniformity lemma, *Geom. Funct. Anal.* 7 (1997), 322-332.

[18] D.J. Kleitman and B.L. Rothschild, Asymptotic enumeration of partial orders on a finite set, *Trans. Amer. Math. Soc.* 205 (1975), 205–220.

[19] Ph.G. Kolaitis, H.J. Prömel and B.L. Rothschild, K_{l+1}-free graphs: asymptotic structure and a 0–1 law, *Trans. Amer. Math. Soc.* 303 (1987), 637–671.

[20] L.H. Loomis and H. Whitney, An inequality related to the isoperimetric inequality, *Bull. Amer. Math. Soc.*, 55 (1949), 961–962.

[21] T. Łuczak, The chromatic number of random graphs, *Combinatorica* 11 (1991), 45–54.

[22] H.J. Prömel and A. Steger, Excluding induced subgraphs: quadrilaterals, *Random Structures and Algorithms* 2 (1991), 55–71.

[23] H.J. Prömel and A. Steger, Excluding induced subgraphs II: extremal graphs, *Discrete Applied Mathematics* 44 (1993), 283–294.

[24] H.J. Prömel and A. Steger, Excluding induced subgraphs III: a general asymptotic, *Random Structures and Algorithms* 3 (1992), 19–31.

[25] H.J. Prömel and A. Steger, The asymptotic structure of H-free graphs, in *Graph Structure Theory* (N. Robertson and P. Seymour, eds), Contemporary Mathematics 147, Amer. Math. Soc., Providence, 1993, pp. 167-178.

[26] F.P. Ramsey, On a problem of formal logic, *Proc. London Math. Soc.* 30(2) (1929), 264–286.

[27] E.R. Scheinerman, Generalized chromatic numbers of random graphs, *SIAM J. Discrete Math.* 5 (1992), 74–80

[28] E.R. Scheinerman and J. Zito, On the size of hereditary classes of graphs, *J. Combinat. Theory (B)* 61 (1994), 16–39.

[29] E. Shamir and J. Spencer, Sharp concentration of the chromatic number on random graphs $G_{n,p}$, *Combinatorica* 7 (1987), 121–129.

[30] E. Szemerédi, Regular partitions of graphs, in *Proc. Colloque Inter. CNRS* (J.-C. Bermond, J.-C. Fournier, M. las Vergnas, D. Sotteau, eds), 1978.

[31] P. Turán, On an extremal problem in graph theory (in Hungarian), *Mat. Fiz. Lapok* 48 (1941), 436–452.

Béla Bollobás
Dept. of Mathematics,
Univ. of Memphis, USA
and
Trinity College,
Univ. of Cambridge,
England