[9]

Applications of Probabilistic Methods to Graph Theory

Paul Erdös

Both in his lectures and his written work, Erdös takes great pleasure in offering a monetary reward to anyone who can settle a stated conjecture; such an offer is made in this lecture. The reader is assured that this offer is bona fide and that Erdös has occasionally had the pleasure of paying. Whenever Erdös spoke (or "preached," as he put it), our seminar audience was always a bit larger than average.

The area of probability in graph theory arose from a theorem of Ramsey, which may be simply explained by the following celebrated problem: Prove that among any six people at a gathering, there will always be three mutual acquaintances or three mutual nonacquaintances.

Here is the basic idea of a probabilistic argument. To prove that there exists a graph with a specific property, one derives an estimate on the number of graphs which do not have the property. If it can be shown that this number is definitely less than the total number of graphs with a given number n of points, then there must exist a graph with the property in question. But this does not give any clue on how to construct such a graph. For example, it had been shown by probabilistic methods that there always exists a graph with an arbitrarily large chromatic number as well as an arbitrarily large girth, but no one had any idea of how to find such graphs. In September 1966, a high school student from Budapest named L. Lován developed a method for constructing them, having nothing to do with the probabilistic proof.

F.H.

The application of probabilistic methods to graph theory stems from a well-known theorem of Ramsey [16], an English philosopher and mathematician, whose brother is the Archbishop of Canterbury. A special case of Ramsey's result may be stated in the language of set theory: For any infinite set S and any partition of $S \times S$ into two subsets T_1 and T_2, there exists an infinite subset A of S such that $A \times A$ is contained in T_1 or in T_2. Clearly the same result holds if the partitioning is into any finite number of subsets.

This result can be stated in graphical language as follows: Every infinite graph contains an infinite complete subgraph or an infinite independent set of points. In other words, for any infinite graph G, G or its complement $\bar{G}$ contains an infinite complete subgraph.

The finite form of Ramsey's theorem is more involved. Let $f(r, s)$ be the smallest integer such that every graph G with $f(r, s)$ points contains at least r mutually adjacent points or s independent points. Symbolically, G contains the complete graph K_r or $\bar{G}$ contains K_s. Of course $f(r, s) = f(s, r)$. An upper bound for the value of $f(r, s)$ was found by Erdös and Szekeres [11], and there has been no serious improvement of this result:

$$f(r, s) \leq \binom{r+s-2}{r-1} \tag{1}$$

The proof of (1) follows readily from the following recursive inequality by double induction on r and s:

$$f(r, s) \leq f(r-1, s) + f(r, s-1). \tag{2}$$

To prove (2), let G be any graph with $f(r-1, s) + f(r, s-1)$ points, and let v be any point of G. Let A be the set of points adjacent with v and B the remaining points. Then A has at least $f(r-1, s)$ points or B has at least $f(r, s-1)$ points, since the sum is the total number of points in G. Assume that A has $f(r-1, s)$ points. Then it follows that the subgraph induced by A contains K_{r-1} or $\bar{K}_s$. If A contains K_{r-1}, then the subgraph induced by v and A contains K_r; otherwise, G has s independent points. If A does not have $f(r-1, s)$ points, then B has $f(r, s-1)$ points, and similar arguments show that G contains K_r or has s independent points, proving inequality (2) and hence the theorem.

The special case of (1) when $r = s$ is

$$f(r, r) \leq \binom{2r-2}{r-1}. \tag{3}$$

This has been improved by Frasnay [12] to the inequality

$$f(r, r) \leq \frac{8}{9}\binom{2r-2}{r-1}. \tag{4}$$

A question on a Putnam examination a few years ago asked the contestants to prove that among any six people at a gathering, there will always be 3 who know each other or 3 who do not know each other. In other words, for any graph G of six points, either G or $\bar{G}$ contains a triangle. Thus this question asks for the proof that $f(3, 3) = 6$.

It is also known that $f(4, 4) = 18$, and it is easily shown that $f(r, 2) = r$. Gleason (unpublished) has recently extended the known values of $f(r, s)$, but the determination of exact values in general remains a difficult unsolved problem, even for $f(r, r)$. The conjecture has been made that $(f(r, r))^{1/r}$ approaches a limit as $r \to \infty$. The following bounds on $(f(r, r))^{1/r}$ are known

$$\sqrt{2} < (f(r, r))^{1/r} < 4\,. \tag{5}$$

The upper bound in (5) follows immediately from (3) since

$$\binom{2r-2}{r-1} < 4^r .$$

The proof by Erdös [3], [6], and [7] of the lower bound uses probabilistic methods as follows. From n given points, fix any r of them and suppose that either all lines or no lines joining them appear, that is, they induce K_r or $\bar{K}_r$. Since there are two choices for each other line of a graph with n points, namely "to be or not to be," there are exactly $2^{\binom{n}{2}-\binom{r}{2}}$ possibilities for the occurrence of these other lines. There are $\binom{n}{r}$ ways of choosing r points from n given points, and there are $2^{\binom{n}{2}}$ labeled graphs with n points. Therefore, for every n satisfying the inequality

$$\binom{n}{r}\cdot 2\cdot 2^{\binom{n}{2}-\binom{r}{2}} < 2^{n/2} , \tag{6}$$

there is a graph with n points containing no K_r or $\bar{K}_r$. Note that the factor 2 in the left side of (6) refers to the choice of K_r or $\bar{K}_r$. But it is easy to verify that (6) holds whenever $n > 2^{r/2}$. Hence $f(r, r) > 2^{r/2}$, completing the proof of (5). Note that no explicit construction is known for such a graph. Only its existence has been proved by this probabilistic (or one might say computational) argument.

An interesting special case is the study of $f(3, s)$. When $r = 3$ is substituted into inequality (1), we get

$$f(3, s) \leq \binom{s+1}{2} . \tag{7}$$

This inequality has not been much improved. But it has been observed by several mathematicians that you can subtract cs, a constant times s, from the right side of (7). Incidentally, a trivial lower bound is given by

$$f(3, s) > 3s. \tag{8}$$

By elementary but complicated computations, Erdös [4] has proved the strict inequality in the following result.

$$\frac{cs^2}{(\log s)^2} < f(3, s) \leq \binom{s+1}{2} - cs . \tag{9}$$

Conjecture. As $s \to \infty$, $f(3, s)/s^2$ approaches a constant. (I offer fifty dollars to anyone who can prove or disprove this conjecture.)

A well-known problem (see Dirac [2]) asked whether there exists a graph G with an arbitrarily high chromatic number containing no triangle. Writing under the pseudonym of Blanche Descartes [1], Tutte proved this result by providing an explicit construction of a graph with an arbitrary chromatic number containing no triangle. As is often the case, this theorem was later rediscovered independently; see Zykov [17] and Mycielski [15].

Tutte's result was extended by Kelly and Kelly [14], who proved by an explicit construction that for any positive integer r, there exists a graph G that contains no triangle, quadrilateral, or pentagon and that has a chromatic number greater than r. They conjectured that, for any two positive integers r and s, there exists a graph G, whose chromatic number is at least r, which contains no polygon with fewer than s sides. This conjecture was proved by Erdös [5], using a probabilistic argument, the outline of which follows.

Consider a graph G with a large number n of points and let c be a large constant. Consider further that G is a random graph with n points and cn lines. In a series of papers on the evolution of random graphs, Erdös and Rényi [10] studied the probable structure of G and its dependence on the value of c. They showed that for sufficiently large c almost all graphs contain triangles, quadrilaterals, and so on. A simple computation shows that the expected number of small cycles is very small. Destroy them by deleting one line from each small cycle. But another simple computation shows that most of the remaining graphs still have an arbitrarily large chromatic number.

Erdös and Hajnal [8] gave a simple construction of an infinite graph having infinite chromatic number and containing no triangle. Let $N^{(2)}$ be the set of all unordered pairs of distinct positive integers $\{i, j\}$ with $i < j$. Now construct a graph G with point set $N^{(2)}$ in which, for every two integers i and k, the pairs $\{i, j\}$ and $\{j, k\}$ are adjacent. Clearly, G has no triangle. To prove that the chromatic number of this infinite graph G is infinite, assume that the chromatic number $\chi(G) = r < \infty$. Then the points of G can be split into r classes of independent points. Applying Ramsey's theorem to the complete graph H whose points are the integers, we find that at least one of these classes contains an infinite complete subgraph H' of H. Therefore, it also contains two lines $\{i, j\}$ and $\{j, k\}$, which shows that G can not have chromatic number r.

A classical problem is to find the maximum number n of points so that the lines of the complete graph K_n can be colored with r colors, in such a way that there is no triangle which is unicolored (all one color). The solution, that $n = 5$ when $r = 2$, is (in disguise) the Putnam examination question mentioned above. Recently, Greenwood and Gleason [13] showed that for 17 points, every coloring with three colors contains a unicolored triangle, but for 16 points, not every coloring contains such. For 66 points and four colors, there always exists a unicolored triangle, but until very recently, the situation for 65 points was unsolved. However, a Hungarian sociologist, Szalai, showed by an explicit construction that one can color the lines of K_{65} by using four colors so that there is no unicolored triangle. Incidentally, in this process Szalai rediscovered Ramsey's theorem.

Erdös and Rado [9] proved the following related theorem concerning infinite graphs without the continuum hypothesis but by using the axiom of choice. For every infinite cardinal number m, there exists a graph G with m points containing no triangle and having chromatic number m.

Erdös conjectured that there exists a graph G containing no quadrilateral and whose chromatic number is nondenumerable. He offered ten pounds to anyone who could prove or disprove it, and Hajnal has just settled this conjecture negatively. Actually, the proof is not difficult.

We conclude by noting that probabilistic methods do not usually give the best possible results, but they can be used in many different situations and enable one to attack problems which could not even be started otherwise.

References

[1] Blanche Descartes, A three colour problem. *Eureka* **9**(April 1947) 21. Solution. *Eureka* **10**(March 1948) 24. (See also the solution to Advanced problem 1526. *Amer. Math. Monthly* **61**(1954) 352.

[2] G. A. Dirac, The structure of k-chromatic graphs. *Fund. Math.* **40**(1953) 42–55.

[3] P. Erdös, Some remarks on the theory of graphs. *Bull. Amer. Math. Soc.* **53**(1947) 292–294.

[4] ———, Graph theory and probability I. *Canad. J. Math.* **11**(1959) 34–38.

[5] ———, Graph theory and probability II. *Canad. J. Math.* **13**(1961) 346–352.

[6] ———, On the number of complete subgraphs contained in certain graphs. *Publ. Math. Inst. Hung. Acad. Sci.* **7**(1962) 459–464.

[7] ———, On a problem in graph theory. *Math. Gaz.* **47**(1963) 220–223.

[8] ——— and A. Hajnal, Some remarks on set theory IX. *Mich. Math. J.* **11**(1964) 107–127.

[9] ——— and R. Rado, A construction of graphs without triangles having pre-assigned order and chromatic number. *J. London Math. Soc.* **35**(1960) 445–448.

[10] ——— and A. Rényi, On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.* **5**(1960) 17–67.

[11] ——— and G. Szekeres, A combinatorial problem in geometry. *Compositio Math.* **2**(1935) 463–470.

[12] C. Frasnay, Sur des fonctions d'entiers se rapportant au théorème de Ramsey. *Comptes Rendus Acad. Sci. Paris.* **256**(1963) 2507–2510.

[13] R. E. Greenwood and A. M. Gleason, Combinatorial relations and chromatic graphs. *Canad. J. Math.* **7**(1955) 1–7.

[14] J. B. Kelly and L. M. Kelly, Paths and circuits in critical graphs. *Amer. J. Math.* **76**(1954) 786–792.

[15] J. Mycielski, Sur le coloriage des graphes. *Colloq. Math.* **3**(1955) 161–162.

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

[17] A. A. Zykov, On some properties of linear complexes. *Math. Sbornik.* (66) **24**(1949) 163–188. *Amer. Math. Soc. Translation No.* 79 (1952).