The Pigeonhole Principle (1)

Ramsey Theory refers to the study of partitions of large structures, and generalizes the pigeonhole principle. A typical result states that a special substructure *must* occur in some class of the partition.

The pigeonhole principle: If a set consisting of more than kn objects is partitioned into n classes, then some class receives more than k objects.

Example. For a simple graph G with 6 vertices, the sum of the degrees of a vertex x in G and $\overline{G}$ is 5, so either $d_G(x) \geq 3$ or $d_{\overline{G}}(x) \geq 3$. (This is the first half of the proof showing that either G or $\overline{G}$ contains a clique of size 3.)

The Pigeonhole Principle (2)

Theorem 8.3.3 Every list with more than n^2 distinct numbers has a monotone sublist of length greater than n.

Proof: Let $a = a_1, a_2, \ldots, a_{n^2+1}$ be the list, and assign position k the label (x_k, y_k), where x_k (y_k) is the length of a longest increasing (decreasing) sublist ending at a_k. With $n^2 + 1$ labels and $1 \leq x_k, y_k \leq n$, at least two of the labels must coincide. But this is not possible: for any $i > j$, a_i increases the length of either of the lists ending at a_j, since $a_i \neq a_j$. □

Ramsey's Theorem (1)

We consider partitions of sets and use the language of coloring—a k-coloring is a partition into k subsets. The set of all r-element subsets (r-*sets*) of S is denoted by $\binom{S}{r}$.

homogeneous A set $T \subseteq S$ is homogeneous under a coloring of $\binom{S}{r}$ if all r-sets in T receive the same color; it is i-*homogeneous* if that color is i.

Ramsey number For given positive integers r and $p_1, p_2, \ldots, p_k$, the smallest integer N such that every k-coloring of $\binom{[N]}{r}$ yields an i-homogeneous set of size p_i for some i; denoted by $R(p_1, p_2, \ldots, p_k; r)$.

Ramsey's Theorem (2)

Example 1. The pigeonhole principle corresponds to $r = 1$.

Example 2. Consider $r = 2$ and $k = 2$. If we view the set S as a set of vertices, then $\binom{S}{r} = \binom{S}{2}$ is the set of edges of a complete graph. These edges are colored with $k = 2$ colors, but since one color class gives the other, we may look at this as a graph G and its complement $\overline{G}$. The Ramsey number is then the smallest integer such that for any graph G of this order, either G has a clique of size p_1 or an independent set of size p_2 (we have seen the case $p_1 = p_2 = 3$ several times earlier).

R. L. Graham, B. L. Rothschild, and J. H. Spencer, *Ramsey Theory*, 2nd ed., Wiley, New York, 1990.

Ramsey's Theorem (3)

Theorem. $R(p_1, p_2; 2) \leq R(p_1 - 1, p_2; 2) + R(p_1, p_2 - 1; 2)$.

Proof: Assuming that $R(p_1 - 1, p_2; 2) = s$ and $R(p_1, p_2 - 1; 2) = t$ exist, let N be their sum. Proving the bound for $R(p_1, p_2; 2)$ means showing that every red/blue-coloring (red=color1, blue=color2) of the edges of a complete N-vertex graph yields a p_1-set of vertices within which all edges are red or a p_2-set of vertices within which all edges are blue.

Consider a red/blue-coloring of K_N, and choose a vertex x. There are $N - 1 = s + t - 1$ vertices other than x, so x has at least s incident red edges or t incident blue edges.

Ramsey's Theorem (4)

Proof: (cont.) By symmetry, we may assume that x has at least s incident red edges. By the definition of s, the complete subgraph induced by the neighbors of x along these edges has a blue p_2-clique or a red $(p_1 - 1)$-clique. The latter would combine with x to form a red p_1-clique. In either case, we obtain an i-homogeneous set of size p_i for some i. □

Since $R(2, p; 2) = R(p, 2; 2) = p$, $R(p_1, p_2; 2)$ is defined for all $p_1, p_2 \geq 2$.

Theorem 8.3.7. Given positive integers r and $p_1, p_2, \ldots, p_k$, there exists an integer N such that every k-coloring of $\binom{[N]}{r}$ yields an i-homogeneous set of size p_i for some i.

Ramsey Numbers (1)

When $r = 2$, we may simply write $R(p_1, p_2, \ldots, p_k)$. Values and best known bounds for $R(p, q)$ with small parameters are as follows:

$p \backslash q$	3	4	5	6	7	8	9
3	6	9	14	18	23	28	36
4		18	25	35–41	49–61	56–84	69–115
5			43–49	58–87	80–143	95–216	121*–316
6				102–165	111–298	127–495	153–780

*: By Harri Haanpää (HUT).

Ramsey Numbers (2)

To find a lower bound on a Ramsey number, $R(p, q) > N$, one should find an explicit red/blue vertex-coloring of an N-vertex complete graph that has neither a red clique of size p nor a blue clique of size q (or prove its existence in some way).

To find an upper bound on a Ramsey number, $R(p, q) \leq N'$ one must show that *every* N'-vertex graph has the desired properties; for example, the recursive theorem $R(p, q) \leq R(p - 1, q) + R(p, q - 1)$ can be used.

Example. $R(3, 3) = 6$. We know that $R(3, 3) \leq 6$. The graph C_5 proves that $R(3, 3) > 5$.

Random Graphs (1)

There are several models for studying random graphs. Making all graphs with vertex set $[n]$ equally likely is equivalent to letting each vertex pair appear as an edge with probability 1/2. This is the most common model for random graphs and leads to the simplest computations. We may allow the probability to depend on n.

Model A. Given n and $p(n)$, generate graphs with vertex set $[n]$ by letting each pair of vertices be an edge with probability p, independently. The random variable G^p denotes a graph generated in this way.

The random graph means Model A with $p = 1/2$.

Random Graphs (2)

We get another model if we fix the number of edges, m.

Model B. Given n and $m(n)$, let each graph with vertex set $[n]$ and m edges occur with the same probability, $\binom{N}{m}^{-1}$, where $N = \binom{n}{2}$. The random variable G^m denotes a graph generated in this way.

It turns out that Model B is accurately described by Model A when n is large and $p = m/\binom{n}{2}$, so one may restrict the attention to Model A.

The Probabilistic Method (1)

The probabilistic method can be used to prove the existence of desired combinatorial objects without constructing them. The main idea: if we take a random object and the probability that it has property P is positive, then there must exist such objects with this property.

N. Alon and J. H. Spencer, *The Probabilistic Method*, 2nd ed., Wiley, New York, 2000.

The Probabilistic Method (2)

Theorem 8.5.4. If $\binom{n}{p}2^{1-\binom{p}{2}} < 1$, then $R(p,p) > n$.

Proof: The bound $R(p,p) > n$ means that there is an n-vertex graph with $\alpha(G) < p$ and $\omega(G) < p$. We use Model A with vertex set $[n]$ and $p = 1/2$. Let Q be the event "neither a p-clique nor an independent p-set".

Each possible p-clique occurs with probability $2^{-\binom{p}{2}}$. The probability of having at least one p-clique is therefore bounded from above by $\binom{n}{p}2^{-\binom{p}{2}}$, and we get the same probability of having at least one independent p-set. Therefore the probability of Q is bounded from below by $1 - 2\binom{n}{p}2^{-\binom{p}{2}}$, and is positive when $\binom{n}{p}2^{1-\binom{p}{2}} < 1$. □

Expectation (1)

There exists an element of the probability space whose value is as large as (or as small as) the expectation $E(X) = \sum_k kP(X = k)$.

Theorem 8.5.8. Some n-vertex tournament has at least $n!/2^{n-1}$ Hamiltonian paths.

Proof: Generate tournaments on $[n]$ randomly by choosing $i \to j$ or $j \to i$ with equal probability for each pair $\{i, j\}$. Let X be the number of Hamiltonian paths; X is the sum of $n!$ indicator variables (taking values 1 or 0 depending on whether we have a Hamiltonian path or not) for the possible Hamiltonian paths. Each Hamiltonian path occurs with probability $1/2^{n-1}$, so $E(X) = n!/2^{n-1}$. □

Expectation (2)

Theorem 8.5.9. $\alpha(G) \geq \sum_{v \in V(G)} \frac{1}{d(v)+1}$.

Proof: Given an ordering of the vertices of G, the set of vertices that appear before all their neighbors form an independent set. When the ordering is chosen uniformly at random, the probability that v appears before all its neighbors is $1/(d(v) + 1)$. Thus the right side of the inequality is the expected size of the independent set formed by choosing the vertices appearing before their neighbors in a random vertex ordering. □

Properties of Almost All Graphs (1)

Given a sequence of probability spaces, let q_n be the probability that property Q holds in the nth space. Property Q **almost always holds** if $\lim_{n\to\infty} q_n = 1$.

Theorem 8.5.18. If p is constant, then almost every G^p has diameter 2 (and hence is connected).

Proof: Let $X(G^p)$ be the number of unordered vertex pairs with no common neighbor. If there are none, then G^p is connected and has diameter 2. By Markov's inequality (if X is nonnegative and integer-valued, then $\lim_{n\to\infty} E(X) \to 0$ implies $\lim_{n\to\infty} P(X = 0) \to 1$ [Wes, Lemma 8.5.17]), we need only show $\lim_{n\to\infty} E(X) \to 0$.

Properties of Almost All Graphs (2)

Proof: (cont.) We express X as the sum of $\binom{n}{2}$ indicator variables, one for each pair $\{v_i, v_j\}$, where $X_{i,j} = 1$ iff v_i and v_j have no common neighbor.

When $X_{i,j} = 1$, the $n - 2$ other vertices fail to have edges to both of these, so $P(X_{i,j} = 1) = (1 - p^2)^{n-2}$ and $E(X) = \binom{n}{2}(1 - p^2)^{n-2}$. When p is fixed, $\lim_{n\to\infty} E(X) \to 0$, and the theorem follows. □