THE PROBABILISTIC METHOD AND RAMSEY'S THEOREM

The probabilistic method is a way to use probability to solve problems that do not seem to have any probabilistic content. It was pioneered by Paul Erdős in the fifties.

Grimmett-Stirzaker has a simple example of the probabilistic method on p. 59. We give another classical example. To state it, we need the notion of a *graph*. A *graph* (or *network*) consists of a set V of vertices and a set E of edges so that between any two distinct vertices, there can either be an edge or not. Formally, $E \subseteq V \times V$ is a symmetric irreflexive relation. That is, it is such that $(x, x) \notin E$ and $(x, y) \in E$ implies $(y, x) \in E$ for all $x, y \in V$.

Consider for example a party with n students, and assume that any pair of students either know or do not know each other (we suppose that knowledge is symmetric: if a knows b, then b knows a). The corresponding graph would have the students as vertices, and an edge between two students if they know each other.

One can ask how big a party should be in order to contain a group of three people that are either all friends or all unacquainted with each other. It turns out that the answer is 6 (prove it!). More generally, we have:

Theorem (Ramsey's theorem)**.** For every natural number r, there is a (very big) natural number n such that any graph with n vertices contains r vertices that are either all connected or all disconnected.

Ramsey's theorem is a result of combinatorics, you do not need to know the proof for this class. Nevertheless we prove it (for completeness) at the end of these notes. Try to prove it for yourself first!

How big should n be in Ramsey's theorem? Erdős used the probabilistic method to show that it must be at least exponential in r:

Theorem. For every natural number $r \geq 3$, there is a graph with at least $\lfloor 2^{r/2-1} \rfloor$ vertices for which any r vertices are neither all connected nor all disconnected.

Proof. We generate a random graph G on n vertices by, for each pair (i, j) of vertices tossing a fair coin and putting an edge between the two if the coin comes up head. That is, with probability $\frac{1}{2}$ there will be an edge between i and j. Now, what is the probability that a fixed subset of r-many vertices is completely connected? Well, there are $\binom{r}{2}$ possible edges, so the probability they are all here is $2^{-\binom{r}{2}}$. Similarly the probability that a fixed subset of r-vertices has has no connections at all is $2^{-\binom{r}{2}}$. There are $\binom{n}{r}$ subsets of vertices of cardinality r, so in total the probability that there are r vertices which are either all connected or all disconnected is bounded by $\binom{n}{r} 2 \cdot 2^{-\binom{r}{2}}$. If this probability is strictly less than 1, this means that there must exist a graph that is as desired. The probability is less than 1 when $\binom{n}{r} < 2^{\binom{r}{2}-1}$. Using that $\binom{n}{r} < n^r$ and $\binom{r}{2} = \frac{r(r-1)}{2}$, we get that this inequality holds when $n \leq 2^{r/2-1}$, as desired. □

Note that the proof is *nonconstructive*: it shows the graph exists but gives no explicit description. In fact, no explicit construction of such a graph is known.

Proof of Ramsey's theorem. Let $R(s,t)$ be the least natural number n such that any graph on n vertices has either a set of vertices of size s that is completely connected or a set of vertices of size t that is completely disconnected. Note that $R(1,t) = R(s,1) = 1$. We prove that for $s,t \geq 2$, $R(s,t) \leq R(s-1,t) + R(s,t-1)$. It then follows (using induction on $s+t$) that $R(s,t)$ is finite and hence that $R(r,r)$ is finite, which is what we wanted. Consider a graph with $R(s-1,t) + R(s,t-1)$ vertices. Let v be a vertex, let A be the set of vertices connected to v and let B be the set of vertices not connected to v. Since $|A|+|B|+1 = R(s-1,t)+R(s,t-1)$, we have that either $|A| \geq R(s-1,t)$ or $|B| \geq R(s,t-1)$. Suppose that $|A| \geq R(s-1,t)$. If the graph has a completely disconnected subset of size t, we are done, so suppose this does not happens. Then by the induction hypothesis A has a completely connected set S of size $s-1$. Then $S \cup \{v\}$ is a completely connected set of size s, as desired. The proof in case $|B| \geq R(s,t-1)$ is similar. $\square$