Randomness and Computation

or, “Randomized Algorithms”

Mary Cryan

School of Informatics
University of Edinburgh

The Probabilistic Method

> *The* Probabilistic Method *is a nonconstructive method of proof, primarily used in combinatorics and pioneered by Paul Erdős, for proving the existence of a desired kind of mathematical object. It works by showing that if we randomly choose objects from a specified class,* the probability *that the result has the desired property is greater than zero. This is enough to tell us that there must be* at least one *object with the desired property in the class.*
>
> *Note that although this approach uses probability, the result (that some object with the property exists) will be definite, not "in probability".*

Slightly different theme to the rest of the results in this course, as we are concerned with showing *existence* (rather than constructing the object). However, sometimes we can derandomize/construct.

Graphs and Colourings

A common concept in graph theory is the concept of a *colouring* of a graph. If we have k different colours, we usually identify them with the set $\{1, \ldots, k\}$.

- We can consider the different ways of colouring the vertices of a graph $G = (V, E)$ with those k colours.
  - A *k-colouring* is any assignment $c : V \to \{1, \ldots, k\}$ of colours to vertices (every $v \in V$ gets some colour $c(v)$).
  - A *proper k-colouring* is any $c : V \to \{1, \ldots, k\}$ such that for every $e = (u, v), e \in E$, we have $c(u) \neq c(v)$.
  - For a given graph $G = (V, E)$, it is often of interest to ask *what is the minimum k* needed to properly colour G. For sure, we know $k \leq$ “max degree of G +1".
  - Lots of research effort have gone into polynomial-time algorithms to approximate (exact is NP-hard) the minimum k for a given G. **Not our concern today**
- Alternatively we can consider the different ways of *colouring the edges* of a graph $G = (V, E)$.

Our example - Ramsay numbers

Our focus today is 2-colouring the **edges** of the complete graph K_n.

- K_n is the *complete graph* on n vertices (for every $i, j \in [n], i \neq j$, we have the edge (i, j)).
- We are **not** interested in vertex 2-colourings of K_n, every vertex “blue" or “red". (cannot give a proper colouring if $n \geq 3$).
- Our concern is whether we can colour K_n's edge with our two colours and make sure that we do not have any “all-blue" or “all-red" subgraph which is “too large".
- The *“Ramsay number"* $R(k, k)$ is the smallest value for n such that in *any two-colouring* of the edges of K_n, there must be **either** be a red K_k (“all-red” of size k) **or** a blue K_k (“all-blue” of size k).

The value of $R(k, k)$ increases with k.

class: What is $R(2, 2)$? And $R(3, 3)$ (board)?

Lower Bound on $R(k, k)$

We prove an *lower bound* on $R(k, k)$ for general k. This was first shown by Erdős in 1947.

Theorem (Theorem 6.1)

Consider $R(k, k)$ for some $k \geq 2$. For any n such that $\binom{n}{k}2^{1-\binom{k}{2}} < 1$, we have $R(k, k) > n$.

Proof.

Write down the *expected* number of “all red"/“all blue" K_k subgraphs, when the edges of K_n are coloured uniformly at random by red/blue.

For a *particular* K_k subgraph, probability of being *monochromatic* is $2 \cdot 2^{-\binom{k}{2}} = 2^{1-\binom{k}{2}}$.

There are $\binom{n}{k}$ different K_k subgraphs to consider in K_n.

The *expected number of monochromatic subgraphs* of K_n is therefore

$$\binom{n}{k}\frac{2}{2^{\binom{k}{2}}}.$$

Lower Bound on $R(k, k)$

Theorem (Theorem 6.1)

Consider $R(k, k)$ for some $k \geq 2$. For any n such that $\binom{n}{k} 2^{1-\binom{k}{2}} < 1$, we have $R(k, k) > n$.

Proof cont'd.

Now if $\binom{n}{k} \frac{2}{2^{\binom{k}{2}}} < 1$ (as per the conditions), this implies that the *expected number of monochromatic K_k subgraphs* is less than 1 when K_n's edges are randomly two-coloured.

Hence there must be *at least* one two-colouring of K_n's edges without any monochromatic K_k subgraph.
So the Ramsay number $R(k, k)$ is larger than any such n.

To be guaranteed a monochromatic K_k we need $\binom{n}{k} \geq 2^{\binom{k}{2}-1}$

□

Lower Bound on $R(k, k)$

Corollary

If $k \geq 3$, then for $R(k, k) > \lfloor 2^{k/2} \rfloor$.

Proof.

Just algebraic manipulation.

Consider $\binom{n}{k} \cdot 2^{1-\binom{k}{2}}$ for the given value of $n = \lfloor 2^{k/2} \rfloor$. This is

$$\begin{aligned}
& \frac{n \ldots (n-k+1)}{k!} \cdot 2^{1-\binom{k}{2}} \\
< \quad & \frac{2^{k/2} \ldots (2^{k/2}-k+1)}{k!} \cdot 2^{1-\frac{k(k-1)}{2}} \\
\leq \quad & \frac{n^k}{k!} \cdot 2^{1+\frac{k}{2}} 2^{-\frac{k \cdot k}{2}} \\
= \quad & \frac{n^k}{2^{\frac{k^2}{2}}} \cdot \frac{2^{1+\frac{k}{2}}}{k!} \\
= \quad & \left(\frac{n}{2^{\frac{k}{2}}} \right)^k \cdot \frac{2^{1+\frac{k}{2}}}{k!} \\
< \quad & 1 \cdot 1,
\end{aligned}$$

as required. $\square$

Making this method *constructive* (“derandomization")

In the proof of Theorem 6.1 about random colourings of K_n and the presence of *any* monochromatic K_ks, we focused on the situation when we have $\binom{n}{k}2^{1-\binom{k}{2}} < 1$. However, the argument shows . . .

Corollary

Let $k \geq 2$. Then for any complete graph K_n, the expected *number of monochromatic K_k subgraphs in a uniform random 2-colouring of the* edges *of K_n is at most $\binom{n}{k}2^{1-\binom{k}{2}}$.*

Corollary

Let $k = 4$. Then for any complete graph K_n, the expected *number of monochromatic K_4 subgraphs in a uniform random 2-colouring of the* edges *of K_n is at most $\binom{n}{4}2^{-5}$.*

Remember the various K_k copies we consider are not necessarily disjoint, expectation is taken over all of them.

Making this method *constructive* (“derandomization")

Using the second Corollary on slide 8, if the expectation is at most $\binom{n}{4}2^{-5}$ (over all random 2-colourings), then *there is some specific 2-colouring* of K_n that has $\leq \binom{n}{4}2^{-5}$ monochromatic K_4 copies.

We can *construct* a specific 2-colouring to satisfy this using the *method of conditional expectation* (and “deferred decisions").

The idea:

- Let f be a specific edge of K_n.
- A random 2-colouring has probability $\frac{1}{2}$ of setting f blue, and probability $\frac{1}{2}$ of setting f red.
- The colours of all the *other* edges are set *uniformly and independently* with probability $\frac{1}{2}$.
- Hence, for at least one of the events $c(f) = \mathrm{red}$, $c(f) = \mathrm{blue}$, the *(conditional) number of expected monochromatic* K_4 is $\leq \binom{n}{4}2^{-5}$.
- Find a way of determining this colour for f, and iterate.

Making this method *constructive* (“derandomization")

Theorem

For every integer n, we can construct *a specific 2-colouring of K_n such that the expected number of monochromatic K_4 subgraphs is at most $\binom{n}{4}2^{-5}$.*

Proof.

To help with the construction, we define a *weight function w* on copies of K_4 which will allow us to measure the *expected* “value" of colouring particular edges blue or red.

Suppose we are part-way through the construction, and some (but not all) edges have their colour fixed.

- We have some partial colouring $c : F \longrightarrow \{\mathrm{blue}, \mathrm{red}\}$, where $F \subseteq E(K_n)$.
- We maintain the invariant that the *expected number of monochromatic K_4 copies*, taken over the *remaining* random 2-colourings for the edges in $E(K_n) \setminus F$, is $\leq \binom{n}{4}2^{-5}$.

Making this method *constructive* (“derandomization")

Theorem

For every integer n, we can construct *a specific 2-colouring of K_n such that the number of monochromatic K_4 subgraphs is at most $\binom{n}{4}2^{-5}$.*

Proof cont'd.

The weight function w assigns a non-negative value to every subgraph K which is a copy of K_4 in K_n. Let $c(K)$ be the set of colours already seen on edges of K, at this stage of the partial colouring. Define

$$w(K) = \begin{cases} 0 & \text{if } c(K) = \{\text{blue}, \text{red}\}. \\ 2^{-5} & \text{if } c(K) = \emptyset \text{ (all edges uncoloured).} \\ 2^{r-6} & \text{if } |c(K)| = 1, \text{ and } r \text{ of } K\text{'s edges have this colour} \end{cases}$$

The *total weight* of the partially coloured K_n is

$$W_F = \sum_{K \text{ a } K_4 \text{ copy in } K_n} w(K).$$

Making this method *constructive* (“derandomization")

Theorem

For every integer n, we can construct *a specific 2-colouring of K_n such that the number of monochromatic K_4 subgraphs is at most $\binom{n}{4}2^{-5}$.*

Proof cont’d.

Note $w(K)$ is the *probability* of that particular K becoming a *monochromatic* copy of K_4 in a *uniform random* 2-colouring of the edges $E(K_n) \setminus F$.

The *expected number of monochromatic K_4 copies* in a *uniform random* 2-colouring of the *so-far uncoloured edges*, is therefore equal to W_F.

To build our “good" 2-colouring, we start with a fixed ordering $e_1, \ldots, e_{n(n-1)/2}$ of the edges of K_n.
$W_\emptyset$ is $\binom{n}{4}2^{-5}$.

Making this method *constructive* (“derandomization")

Proof.

THE ALGORITHM:

1. **for** $i \leftarrow 1$ **to** $n(n-1)/2$ **do**

   (F is $e_1, \ldots, e_{i-1}$, and these edges are coloured)

2. Calculate W_{red}, the effect on W_F of colouring e_i red.
3. Calculate W_{blue}, the effect on W_F of colouring e_i blue.
4. **if** $W_{\text{red}} < W_{\text{blue}}$ **then** Set $c(e_i) = \text{red}$; $W_F \leftarrow W_{\text{red}}$
5. **else** Set $c(e_i) = \text{blue}$; $W_F \leftarrow W_{\text{blue}}$
6. $F \leftarrow F \cup \{e_i\}$

- The value of W_F *never increases* through the iterations. Hence we end up with a colouring c with at most $W_\emptyset = \binom{n}{4}2^{-5}$ monochromatic K_4s.
- e_i can belong to at most n^2 K_4s in K, so the W_{red}, W_{blue} values can be calculated in $\Theta(n^2)$ time.

The Probabilistic Method in Derandomization

- The theorem on slides 10-13 can be considered to be a “derandomization" of the result on *expected number of monochromatic K_4s in K_n*.
  - We were able to use *conditional expectation* to construct a specific colouring with less than or equal to the expected number of monochromatics.
  - Our algorithm was in fact *polynomial-time* (about n^2 iterations, each doing $\Theta(n^2)$ work, so roughly $\Theta(n^4)$).
- We previously used conditional expectation to derandomize the MAX-CUT algorithm in lecture 6.

Reading and Doing

Reading

- You will want to read Sections 6.1, 6.2, 6.3 from the book.
- The Theorem on derandomizing monochromatic K_4 is not in the book.