Randomness and Computation
or, “Randomized Algorithms”

Mary Cryan

School of Informatics
University of Edinburgh
Making the $K_n$ 2-edge-colouring constructive

**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.**

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

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

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")

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 cont’d.

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 $W_{red} < W_{blue}$ then Set $c(e_i) = \text{red}; W_F \leftarrow W_{red}$

5. else Set $c(e_i) = \text{blue}; W_F \leftarrow W_{blue}$

6. $F \leftarrow F \cup \{e_i\}$

Note that the value of $W_F$ never increases through the iteration of this process. Hence we end up with a colouring $c$ which has fewer than $W_{\emptyset} = \binom{n}{4}2^{-5}$ monochromatic $K_4$s.

$e_i$ can belong to at most $n^2$ $K_4$s in $K$, so the $W_{red}$, $W_{blue}$ values can be calculated in $\Theta(n^2)$ time.
The theorem on slides 2-5 can be considered to be a “derandomization” of the result on expected number of monochromatic $K_4$s 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 can use the method of conditional probabilities to derandomize algorithms like the MAX-CUT algorithm from lectures 3-4.
new topic: The Probabilistic Method cont’d.

In Lecture 10 (and continuing today) we saw the application of the Probabilistic Method:

- To allow us to set bounds on certain parameters that will ensure a randomly-drawn combinatorial object (from whatever pool of possible objects we are focusing on) has some desirable property with probability $> 0$.
  
  (our example property was that the edge 2-colouring of $K_n$ graph would be without any monochromatic $K_k$ subgraphs, assuming $n$ is large enough wrt a lower bound)

  - The “probabilistic method” then allows us to infer that at least one of the combinatorial objects (from our pool) must have the desired property.

- Sometimes we can also derandomize this existence proof and actually construct an object satisfying the desired property.
  
  (need to have a de-composable (wrt deferred decisions) for drawing the random object, then apply conditional probabilities)
Previous slide refers to “having some desirable property with probability > 0”.

In practice, often the approach will be to evaluate with an expectation rather than a probability.

Many of the examples of the probabilistic method we meet in RC involve showing that we can construct a combinatorial object (from some pool) that avoids having some banned sub-structure.

- We can consider the expected number of the banned substructures, when we draw an object from the sample pool;

- Sometimes it will be possible to evaluate the expected number of banned-substructures. If this is < 1, then the probability that there are some combinatorial objects that avoid all banned-substructures is > 0.
Examples of the probabilistic method so far only worked with *expectation*. If we had $E[\cdot] < 1$ for the number of banned sub-structures, we knew there must be some object that has none of the banned sub-structures at all.

However, if we have $E[\cdot] > 1$, things are less clear. There are definitely objects containing the banned sub-structures, but how *likely* they are is not clear.

Early on in the course we gave the definition of the *second moment* of a discrete random variable $X$, this being $E[X^2]$. Then variance is $E[X^2] - E[X]^2$.

The second moment (with Chebyshev) can help us show that a typical sample is likely to have $X$ close to $E[X]$.  

*RC (2017/18) – Lecture 11 – slide 8*
Second Moment Method

Theorem (Theorem 6.7)
For any integer-valued random variable $X$ with positive expectation, we have

$$\Pr[X = 0] \leq \frac{\text{Var}[X]}{(\text{E}[X])^2}.$$ 

Proof.
Really just a special case of Chebyshev's Inequality. We are interested in $\Pr[X = 0]$, which is equal to $\Pr[\text{E}[X] - X = \text{E}[X]]$. Also,

$$\Pr[\text{E}[X] - X = \text{E}[X]] \leq \Pr[\text{E}[X] - X \geq \text{E}[X]] \leq \Pr[|\text{E}[X] - X| \geq \text{E}[X]].$$

Then this final $\Pr[\cdot]$ fits the form for Chebyshev’s Inequality with $a = \text{E}[X]$, so applying Chebyshev gives us

$$\Pr[|\text{E}[X] - X| \geq \text{E}[X]] \leq \frac{\text{Var}[X]}{(\text{E}[X])^2},$$

and this right-hand side also bounds $\Pr[X = 0]$. 

RC (2017/18) – Lecture 11 – slide 9
We are interested in the random model $G_{n,p}$, where we draw a random graph on $n$ vertices by independently doing a Bernoulli trial for each potential edge $(u, v)$, $u \in V$, $v \in V \setminus \{u\}$, adding the edge with probability $p$, omitting that edge if the trial returns 0.

We are interested in whether the drawn graph $G \leftarrow G_{n,p}$ contains a 4-clique or not.

Clearly the graph is more likely to have a 4-clique if $p$ has a higher value (and $G \leftarrow G_{n,p}$ is likely to have more edges).

In lecture 12, we will show that there is a *threshold* for “$G \leftarrow G_{n,p}$ having a 4-clique” when $p$ is either side of $\Theta(n^{-2/3})$. 

**Threshold for 4-cliques in $G_{n,p}$**
Changes to this slide set

The original slides planned for 9th March had material on the 2nd moment method and its application to analysing the likelihood of a 4-clique in graphs from $G_{n,p}$. We got slowed down by de-randomization of the process for generating a “good” 2 edge-colouring, and discussion of administrative issues. I have moved the material on 4-cliques in $G_{n,p}$ to lecture 12.

There were a number of typos in the derandomization of the “good” (no monochromatic $K_4$) 2-edge-colouring process, which have now been fixed.

- Use of $K_k$ instead of $K_4$ sometimes.
- Missing detail about updates to $W_F$ and $F$ now added to the derandomization algorithm.
- Under the algorithm, have corrected the claim about $W_F$ (it never increases).
Reading and Doing

Reading

▶ The Theorem on derandomizing monochromatic $K_4$ is not in the book.

▶ You will want to read Sections 6.2, 6.3, 6.4, 6.5, 6.6 from the book.

▶ As preparation for Lecture 12, you should read 6.7.

Course admin

▶ There will be a tutorial in week 8 (Wednesday 12 in the usual spot). I will send out a tutorial sheet for next week (week 8) over the weekend.

▶ I will probably upload Lecture 12 to the website over the coming weekend.