Graph colouring via the probabilistic method

Bruce Reed
Équipe Combinatoire CNRS
Université Pierre et Marie Curie, Paris, France

April 21, 1997
[summary by F. Morain and P. Robert]

1. Introduction

Colouring a graph with the minimum number of colours is a classical problem in graph theory and has many applications. For instance, think of a cellular phone network on which each vertex (phone) must use a different frequency with its neighbours. This problem is also known to be a difficult one (see for instance [3]).

The purpose of the talk is to present a naive algorithm for colouring a certain type of graphs and explain how to analyze it with elementary probabilistic tools that we will describe first.

2. Probabilistic tools

Throughout this section we denote by \( \text{Pr}(A) \) the probability of the event \( A \), \( E(X) \) the expected value of the random variable \( X \), and \( E(X/A_1, \ldots, A_n) \) the conditional expectation of \( X \) relative to the events \( A_1, \ldots, A_n \).

2.1. The Lovász Local lemma. Suppose that on some probability space \( \Omega \), there are \( n \) events \( A_1, \ldots, A_n \) that are undesirable. We wish to estimate if there is a positive probability to avoid any of them, i.e., if there is a positive lower bound for the quantity

\[
\Delta = \text{Pr}(\cap_{i=1}^n A_i^c),
\]

where \( A_i^c = \Omega - A_i \). If the events are independent, that is for any \( k \)-tuple \( 1 \leq i_1 < \cdots < i_k \leq n \),

\[
\text{Pr}(\cap_{j=1}^k A_{i_j}) = \prod_{j=1}^k \text{Pr}(A_{i_j}),
\]

then

\[
\Delta = \prod_{i=1}^n (1 - \text{Pr}(A_i)).
\]

The problem is that in practice, the events are not always completely independent but weakly independent, in the sense that for each \( i \) there exists a subset \( V_i \subset \{1, \ldots, n\} \) such that \( A_i \) is independent of the events \( A_j, j \in V_i^c \). In other words, \( A_i \) is possibly dependent of \( A_j \) with \( j \) in the “neighbourhood” \( V_i \) of \( i \). If the cardinality of the \( V_i \)’s is small, one might expect an estimate close to the one we saw for the independent case. This is the conclusion of Lovács’s lemma, see [1].

Lemma 1. If the events are such that for all \( 1 \leq i \leq n \),
1. \( \Pr(A_i) \leq p \),
2. \( A_i \) is independent of \( (A_j)_{j \neq i} \),
3. \( |V_i| \leq d \),

and if \( p(d+1) < 1 \) then none of the events \( A_i, i = 1, \ldots, n \), occurs with positive probability.

2.2. Azuma’s Inequality. If \( (Y_i) \) be a sequence of independent random variables with the same distribution on \( \{0, 1\} \), \( p = \Pr(Y_i = 1) \); the result of successive coin tossings is a good model for this sequence of random variables. It is well known that the time averages \( \frac{1}{n} \sum_{i=1}^{n} Y_i \) converges exponentially fast to \( p \) as \( n \to +\infty \). Rigourously, this is Chernoff’s bound

\[
\Pr \left( \left| \frac{1}{n} \sum_{i=1}^{n} Y_i - p \right| > a \right) < 2e^{-a^2/3np},
\]

it says basically that with high probability \( [p/a^2] \) coin tossings are sufficient to get an estimate of \( p \) with an accuracy of the order of \( a \). This kind of result has been extended, for independent variables, to the case of arbitrary distributions, i.e., not only with values in \( \{0, 1\} \), as long as they have an exponential moment. This is a part of large deviations theory, see [2].

Another possible generalization is to consider the case where instead of the sum of independent variables, one looks at some functional \( X \) of some arbitrary random variables \( Y_1, \ldots, Y_n \) with values in \( \{0, 1\} \). Azuma’s inequality says that if the conditional exceptions of \( X \) with respect to \( Y_1, \ldots, Y_i \) do not jump sharply as \( i \) goes from 1 to \( n \), then \( X \) is concentrated around its average value, formally,

**Proposition 1.** If for each \( i \leq n \),

\[
\max_{y_1, \ldots, y_i+1 \in \{0,1\}} |E(X/Y_1 = y_1, \ldots, Y_i = y_i, Y_{i+1} = y_{i+1}) - E(X/Y_1 = y_1, \ldots, Y_i = y_i)| \leq c_i,
\]

then

\[
\Pr (|X - E(X)| > a) \leq 2e^{-\frac{a^2}{2\sum_{i=1}^{n} c_i^2}}.
\]

Azuma’s inequality is surprisingly sharp considering the weak hypotheses of the proposition. In the independent case, for \( X = \sum_{i=1}^{n} Y_i \), condition (1) is satisfied with \( c_i = 1 \), hence the inequality is in this case,

\[
\Pr \left( \left| \frac{1}{n} \sum_{i=1}^{n} Y_i - p \right| > a \right) < 2e^{-\frac{a^2}{3np}},
\]

which is very close to Chernoff’s bound.

3. Graph Colouring

We colour a graph \( G \) such that every pair of adjacent vertices receive different colours. The chromatic number of \( G \), noted \( \chi(G) \), is the minimum number of colours required to colour \( G \). It is easy to see that, if \( \Delta(G) \) denotes the maximal degree of \( G \), then \( \chi(G) \leq \Delta(G) + 1 \).

We can obtain good bounds for \( \chi(G) \) for certain types of graphs, as explained in [4]. For fixed \( \varepsilon > 0 \), we saw that a vertex \( v \) is \( \varepsilon \)-sparse if the subgraph induced by \( N_v \), the neighbourhood of \( v \), has at most \((1 - \varepsilon) \frac{\Delta(\Delta-1)}{2}\) edges. A graph is \( \varepsilon \)-sparse if each of its vertices is \( \varepsilon \)-sparse.

**Theorem 1.** For \( \Delta \) sufficiently large, if \( G \) has maximum degree \( \Delta \) and \( G \) is \( \varepsilon \)-sparse, then \( \chi(G) \leq (1 - \varepsilon/2e^6) \Delta \).
Let us indicate a rough proof of this theorem. In a first step, we construct a partial colouring \( C \) of \( G \) such for each vertex \( v \), the number of neighbours of \( v \) which are coloured exceeds the number of colours appearing on \( N_v \) by at least \( \frac{\varepsilon}{2e\Delta} \Delta + 1 \).

From this, we complete the colouring of \( C \) to a \((1 - \frac{\varepsilon}{2e\Delta}) \Delta\)-colouring of \( G \) in a greedy manner: We colour the remaining vertices one at a time. When we come to colour \( v \), there must be an available colour: Since \( v \) has at most \( \Delta \) neighbours (this is where the sparseness comes in), the number of colours appearing in \( N_v \) is bounded by

\[
\Delta - \left( \frac{\varepsilon}{2e\Delta} \Delta + 1 \right).
\]

Hence fewer than \((1 - \frac{\varepsilon}{2e\Delta}) \Delta\) colours appear in its neighbourhood.

Let us come back to the construction of \( C \). We first assign each vertex of \( G \) a uniformly random colour from \( \{1, 2, \ldots, \lceil \Delta/2 \rceil \} \). If two adjacent vertices have the same colour, we uncolour them. The resulting partial colouring yields \( C \).

The first thing to show is that \( C \) is not too small, which is rather easy. Then we must study, for vertex \( v \), the random variable \( Z_v \) which counts the number of pairs of vertices in \( N_v \) which have the same colour in \( C \). It can be shown that since \( G \) is sparse, the expectation of \( Z_v \) is greater than \( \varepsilon \Delta/e^4 \).

Now that we have proved that many vertices in \( N_v \) are coloured, we must show that \( Z_v \) does not differ too much from its expected value. Once this is done, we use the Local Lemma to prove that every vertex will have such a property, thus proving the property on \( C \). By a technical argument replacing \( Z_v \) with a more amenable quantity, Azuma’s Inequality is used to prove the assumption on \( Z_v \). Roughly speaking, the idea is that a colouring of \( v \) should not influence the colouring of the other parts of \( C \), since \( G \) is sparse.

References


