Probability and Algorithms — Caltech CS150, Winter 2003

Leonard J. Schulman — Scribe: Wonjin Jang

Notes for lecture 12, Feb 24, 2003. Sampling graph colorings.

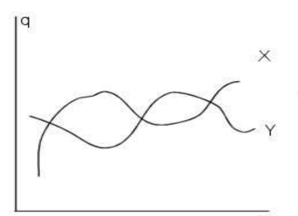
Let n be the number of vertices and Δ be the max degree in G. We wish to sample q-colorings uniformly. For the decision problem, the following is known - If $q < \Delta$, it's a computationally hard problem.
If $q > \Delta$, there always exists a solution.
If $q = \Delta$, there is a polynomial time algorithm to solve the decision problem. (There are just a few special graphs which cannot be q-colored, and we can recognise them.)

For the sampling problem (sample approximately uniformly at random among all colorings), the status is that if $q \geq \frac{11}{6}\Delta$, there is a poly-time sampling method.

As an aside, it is known that the number of colorings of a graph can be evaluated exactly by evaluating a certain polynomial called the *chromatic polynomial* of the graph. The above stated facts imply that this polynomial can be (computationally feasibly) closely approximated for $q \geq \frac{11}{6}\Delta$, but not for $q < \Delta$ unless P=NP. The cases in between are open.

1 Random sampling of graph colorings with $q \geq 4\Delta + 1$

We'll show poly-time approximately uniform sampling in the case that $q \geq 4\Delta + 1$ (proving the result for $q \geq \frac{11}{6}\Delta$ is outside the scope of this class) and prove it by a coupling on q-colorings of G.
The Markov Chain is as follows:
Pick $v \in V, c \in C$ uniformly. Recolor v to c unless some neighbour of v is already colored with c.
This Markov chain is
1. Non-periodic,
2. Connected.
These conditions imply ergodicity of the chain.



Coupling: Let X and Y be two colorings of the graph, represented graphically in the above figure. Let V_1 be the set of vertices in which X agrees in color with Y. Define the distance between two colorings as

$$d \stackrel{\text{def}}{=} |V| - |V_1|.$$

The coupling process is straightforward: choose a vertex v and a color c uniformly and randomly, and recolor v to c in each coloring if it is possible.

Call the step "good" if the recoloring can be done in both chains, and if X and Y had not agreed at v prior to the recoloring. Call the step "bad" if the recoloring can be done in just one chain, but X and Y had agreed at v prior to the recoloring. (Note that a step may be neither good nor bad.)

Run the above Markov chain on the initial coloring X_1 and Y_1, to successively obtain the colorings X_t and Y_t, until the two colorings X_t and Y_t are identical.

$$\begin{aligned} d_t &\stackrel{\text{def}}{=} |V - V_1| \\ d_{t+1} &= d_t - 1 \text{ if the last step was "good"} \\ d_{t+1} &= d_t + 1 \text{ if the last step was "bad"} \\ d_{t+1} &= d_t \text{ otherwise} \end{aligned}$$

The total number of moves available to the algorithm in each step is nq.

The number of "good" moves is $\geq d_t(q - 2\Delta)$, since there are at most Δ neighbours for each $v \in V - V_1$, and each of these neighbors has one color in X_t, and another (possibly equal) in Y_t, with which v cannot be recolored.

The number of "bad" moves is $\leq 2d_t\Delta$ (as above) because a bad move is an attempt to recolor a neighbor v of a vertex $w \in V - V_1$, with one of the colors $X_t(w)$ or $Y_t(w)$.

$$\begin{aligned} E(d_{t+1}|d_t) &\leq d_t(1 + \frac{2\Delta}{qn} - \frac{q-2\Delta}{qn}) \\ &= d_t(1 - \frac{q-4\Delta}{qn}) \\ E(d_t) &\leq n(1 - \frac{q-4\Delta}{qn})^t \\ P(d_t > 0) &\leq n(1 - \frac{q-4\Delta}{qn})^t \\ E(\text{coupling time}) &\in O(\frac{qn}{q-4\Delta}\log n) \end{aligned}$$

This method can be extended to $q \geq 2\Delta + 1$ with a bit more work. $11\Delta/6$ is harder.

2 Approximatly counting the number of colorings

Now that we know how to sample approximately uniformly from the colorings of a graph, subject to $q \geq 4\Delta + 1$, let's see how to count those colorings. (This doesn't follow automatically because the coloring problem isn't self-reducible.)

We are given a graph G with m edges and maximum degree Δ, and a parameter $q \geq 4\Delta + 1$. Order the edges arbitrarily $e_1 \ldots e_m$. Let G_i be the subgraph of G with edges $e_1 \ldots e_i$. Let $\Omega(G_i) = \{q\text{-colorings of } G_i\}$. Then we have the following relations.

$$\begin{aligned} \Omega(G_{i+1}) &\subseteq \Omega(G_i) \\ \Omega(G_0) &= \{1, ..., q\}^n \\ \Omega(G_m) &= \Omega(G) \end{aligned}$$

We will estimate $|\Omega(G)|$ by estimating each ratio $\frac{|\Omega(G_{i+1})|}{|\Omega(G_i)|}$, and multiplying out.

$$|\Omega(G)| = |\Omega(G_0)| \prod_{i=0}^{m-1} \frac{|\Omega(G_{i+1})|}{|\Omega(G_i)|} = q^n \prod_{i=0}^{m-1} \frac{|\Omega(G_{i+1})|}{|\Omega(G_i)|}$$

We will estimate each of the ratios $\frac{|\Omega(G_{i+1})|}{|\Omega(G_i)|}$ by sampling nearly uniformly from $\Omega(G_i)$ and checking whether the coloring is in $\Omega(G_{i+1})$. This will give us a reliable multiplicative estimate of the ratio because of:

Lemma: If $\Delta \geq 2$ then

$$\frac{|\Omega(G_{i+1})|}{|\Omega(G_i)|} \geq 7/8.$$

(If $\Delta < 2$ it is trivial to sample or count colorings of the graph.)

Proof:- Consider any coloring X in $\Omega(G_i) - \Omega(G_{i+1})$. It assigns the same color to the two vertices on e_i. There are at least $q - \Delta$ ways to recolor the higher-indexed of these two vertices to form a legal coloring of G_{i+1}, and the original coloring can be reconstructed from the modified one simply by assigning to the higher-indexed vertex the color of the lower-indexed one. Therefore

$$\begin{aligned} |\Omega(G_{i+1})| &\geq (q-\Delta)(|\Omega(G_i) - \Omega(G_{i+1})|) \\ &= (q-\Delta)(|\Omega(G_i)| - |\Omega(G_{i+1})|) \end{aligned}$$

$$\begin{aligned} \frac{|\Omega(G_{i+1})|}{|\Omega(G_i)|} &\geq \frac{q-\Delta}{q-\Delta+1} \\ &\geq \frac{3\Delta+1}{3\Delta+2} \\ &\geq 7/8 \quad \text{since } \Delta \geq 2. \end{aligned}$$