CS 49/249: Randomized Algorithms (Spring 2021) : Reading&Writing

Topic: Decentralized Graph Coloring from a paper by Bertschinger et al. [BLM+20]

Disclaimer: These notes have gone through scrutiny, but they still probably have errors.
Please send any errors you find to ankita.sarkar.gr@dartmouth.edu

---

Other than the cited papers, the following resources were useful in preparing this note:

1. DeepC's lecture note template and writing guidelines.
2. These writing tips, especially the "Citations and References" section.
3. Online TEX tools – Overleaf, Detexify, TeX StackExchange.
4. Wikipedia articles on probability theory.

Content begins on the following page. Enjoy!

- ***Graph coloring.*** Here is a classic problem: consider a graph $G = (V, E)$ and a palette of colors P. If each vertex is painted with a color from P, we get a *coloring*, $c : V \to P$. We call a coloring *proper* if no edges are *monochromatic*, i.e. no $uv \in E$ should have $c(u) = c(v)$. A proper coloring is easy to get if $|P| = |V|$; but in the graph coloring problem, we have to use the fewest possible colors.

> **Remark:** *An application of graph coloring is in channel selection for WiFi routers. Routers that are close to each other must use different frequency channels to avoid interference. The number of available channels is far smaller than the total number of routers.*

Graph coloring is NP-hard, so it is worthwhile to look at useful special cases. For instance, if we knew that α colors suffice for a proper coloring, could we produce such a coloring?

- α can be $(\Delta + 1)$, where Δ is the maximum degree in G. A simple induction argument shows that $(\Delta+1)$ colors will suffice to produce a proper coloring. This is also tight, e.g. in complete graphs. So $\alpha = (\Delta + 1)$ is an important special case. We consider the problem of producing a $(\Delta + 1)$-coloring, i.e. a proper coloring that uses at most $(\Delta + 1)$ many colors.

> $(\Delta + 1)$-COLORING
>
> **Input.** A graph $G = (V, E)$, $|V| = n$, $|E| = m$, $\max_{v \in V} \deg_G(v) = \Delta$.
>
> **Output.** $c : V \to [\Delta + 1]$ such that $\forall uv \in E,\ c(u) \neq c(v)$.

Here is a simple randomized approach: for each vertex v, pick $c(v) \in [\Delta + 1]$ independently and uniformly at random. Then, each edge has probability $1/(\Delta + 1)$ of being monochromatic, so the expected number of monochromatic edges is $m/(\Delta + 1)$. This can be $\Theta(n)$, e.g. in complete graphs – so we try to "fix" the monochromatic edges by randomly recoloring some vertices.

- ***The algorithm.*** After the first random coloring, if there are vertices with neighbors of the same color, then pick one such vertex v at random. Repick $c(v)$ uniformly at random from $[\Delta + 1]$. If there still are monochromatic edges, then repeat.

```
1: procedure RandColor(G(V, E), Δ)
2:     ▷ c will be our coloring, c : V → [Δ + 1]
3:     for v ∈ V do   ▷ initial random coloring
4:         c(v) ←Random([Δ + 1])
5:     while ∃ uw ∈ E, c(u) = c(w) do   ▷ as long as there is a monochromatic edge
6:         v ←Random({v | ∃u ∈ N(v), c(u) = c(v)})   ▷ pick random "violating" vertex
7:         c(v) ←Random([Δ + 1])   ▷ random recoloring of v
8:     return c
```

RANDCOLOR runs until it produces a proper coloring, so we want to know the *expected* number of recoloring steps, i.e. iterations of Line 5, that are needed to get to a proper coloring.

- ***Intuition.*** RANDCOLOR only recolors one vertex at a time. Let c_t be the coloring after the t^{th} recoloring step; if c_{t-1} assigned the color pink to 5 vertices, then at least 4 vertices remain pink in c_t. Also if, at some step, there is exactly one pink vertex, then that vertex will never be recolored until it gets a pink neighbor. So once there is pink in the graph, the number of pink vertices never drops below 1. Hence, consider this intermediate goal: get all the $(\Delta + 1)$ colors *somewhere* in the graph.

In a complete graph, this intermediate goal is equivalent to the final goal – if we have n colors in a graph, it must be properly colored. So RANDCOLOR behaves as follows on a complete graph: each recoloring step introduces a new color into G with some probability, and once we have seen all n colors, we are done. But this is a familiar problem – *Coupon Collector*! Thus we can expect that, on a complete graph, $O(n \log n)$ recolorings will produce a proper coloring.

We extend this intuition to general graphs as follows: for a vertex v, if $\{v\} \cup N(v)$ has $(\deg_G(v) + 1)$ colors in it, then it is properly colored. So "locally" in $\{v\} \cup N(v)$, expected $O(\Delta \log \Delta)$ recoloring steps give a proper coloring. Crudely, each of these sets have $\approx \Delta$ vertices, so V is made up of $\approx n/\Delta$ such sets. Hence, we guess that RANDCOLOR takes expected $O((n/\Delta) \cdot \Delta \log \Delta) = O(n \log \Delta)$ recoloring steps to get a proper coloring globally. This guess is, happily, correct [BLM+20].

- ***Analysis.*** We will employ a familiar tactic. We define a potential function that decreases as we get "closer" to a proper coloring, and is zero at a proper coloring. A natural candidate for this function is the size of M_t, the set of monochromatic edges after the t^{th} recoloring step.

  Let $\mathbf{\Phi}_M(t) := |M_t|$. Qualitatively, if each recoloring step caused an expected δ decrease in $\mathbf{\Phi}_M$, then we would need an expected $\mathbf{\Phi}_M(0)/\delta$ recoloring steps. This can be formalized via a theorem due to He and Yao [HY04], which gives the following when applied to our $\mathbf{\Phi}_M$.

  Theorem 1 (Additive Drift of $\mathbf{\Phi}_M$)**.** *T be the first t where $\mathbf{\Phi}_M(t) = 0$. If $\exists\, \delta > 0$ such that $\forall s \in [m]$ and $\forall t > 0$ we had*

  $$\mathbf{Exp}[(\mathbf{\Phi}_M(t) - \mathbf{\Phi}_M(t-1)) \;\mid\; \mathbf{\Phi}_M(t-1) = s] \leq -\delta$$

  then we have

  $$\mathbf{Exp}[T] \leq \mathbf{Exp}[\mathbf{\Phi}_M(0)] \;/\; \delta$$

  We know that $\mathbf{Exp}[|M_0|] = m/(\Delta + 1)$. We need $\mathbf{Exp}[(|M_t| - |M_{t-1}|) \;\mid\; |M_{t-1}|]$. If v is recolored at the t^{th} step, then c_{t-1} and c_t can differ only at v. If $c_{t-1}(v)$ was pink, and v was picked for recoloring, then v must have pink neighbors. Let C be the set of vertices that are reachable from v via only pink vertices; we will call C a *maximal monochromatic component* containing v. Then, exactly $\deg_C(v)$ neighbors of v are pink. With probability $\Delta/(\Delta + 1)$, $c_t(v) \neq c_{t-1}(v)$, so the $\deg_C(v)$ pink neighbors of v will no longer form monochromatic edges with v. But, for each non-pink neighbor u of v, $c_t(v)$ could be $c_t(u) = c_{t-1}(u)$ with probability $1/(\Delta+1)$; this creates the new monochromatic edge uv. Hence, conditioned on the event $\mathcal{E}_v$ that v is recolored at the t^{th} step,

  $$\mathbf{Exp}[(|M_t| - |M_{t-1}|) \;\mid\; |M_{t-1}|, \mathcal{E}_v] = \frac{\Delta}{\Delta+1} \cdot (-\deg_C(v)) + \frac{1}{\Delta+1} \cdot (\deg_G(v) - \deg_C(v)) \tag{1}$$

  $$\leq -\deg_C(v) + \frac{\Delta}{\Delta+1} \tag{2}$$

  If we condition, instead, on the event $\mathcal{E}_C$ that *some* vertex $v \in C$ gets recolored, then we get

  $$\mathbf{Exp}[(|M_t| - |M_{t-1}|) \;\mid\; |M_{t-1}|, \mathcal{E}_C] \leq -\overline{d}(C) + \frac{\Delta}{\Delta+1} \tag{3}$$

where $\overline{d}(C)$ is the average $\deg_C(w)$ over all $w \in C$. When $|V(C)| \geq 3$, we have[1] $\overline{d}(C) \geq 4/3$, so the RHS in Equation (3) becomes $\leq -1/3$. Then in Theorem 1, we could put $\delta = 1/3$, to get $\mathbf{Exp}[T] \leq 3m/(\Delta+1)$. That, by the handshake lemma, is $O(n)$.

Sadly, we know that $|V(C)| = 2$ will occur. As we get closer to a proper coloring, we expect C's to be smaller, so we will see *isolated monochromatic edges*, i.e. C's of size 2. Then, we can only say that $\overline{d}(C) \geq 1$, so $\delta = 1/(\Delta+1)$, which only assures an expected $m = O(n\Delta)$ recolorings.

- ***Towards a better potential function.*** Motivated by the above, we tweak our potential function to include I_t, the set of isolated monochromatic edges after the t^{th} recoloring step. Let $\mathbf{\Phi}_I = \Theta(|I_t|)$. We will fix constants suitably later, to ensure that $(\mathbf{\Phi}_M + \mathbf{\Phi}_I)$ behaves the way we want. Now, let us study the effect of recolorings on $|I_t|$.

When a pink v is recolored to green, each of its $\deg_C(v)$ pink neighbors could form one new isolated pink edge of the form uw, $u \in N(v)$, $w \in N(u)$. Also, if v has a green neighbor x that was *properly colored* in c_{t-1}, i.e. all of $N(x)$ was non-green in c_{t-1}, then vx would become a new isolated green edge. So up to $(\deg_C(v)+1)$ new isolated monochromatic edges could be created, giving us

$$\mathbf{Exp}[(|I_t| - |I_{t-1}|) \;\mid\; |I_{t-1}|, \mathcal{E}_C] \leq \overline{d}(C) + 1 \tag{4}$$

The above is an increase; but the hope is that when $\mathbf{\Phi}_M$ is big, it will dominate over $\mathbf{\Phi}_I$, and when $\mathbf{\Phi}_M$ is small, the above bound can be improved. Indeed, when $\mathbf{\Phi}_M$ is small, most C's affected by recoloring will be isolated monochromatic edges. When $C = uw$ such that $c_{t-1}(u) = c_{t-1}(w) =$ pink, then conditioned on $\mathcal{E}_C$, each of u or w gets recolored with probability $1/2$. Wlog, say u gets recolored. With probability $\Delta/(\Delta+1)$, u gets a non-pink color, perhaps green. Then uw is no longer monochromatic, but u could form new isolated green edges. Where P_t is the set of properly colored vertices after step t, it is precisely u's neighbors in P_{t-1} that could form a new isolated green edge with u. This occurs with probability $1/(\Delta+1)$ per vertex in $N(u) \cap P_{t-1}$. So we get

$$\mathbf{Exp}[(|I_t| - |I_{t-1}|) \;\mid\; |I_{t-1}|, \mathcal{E}_C, C = uw] \leq -\frac{\Delta}{\Delta+1} + \frac{|N(u) \cap P_{t-1}| + |N(w) \cap P_{t-1}|}{2(\Delta+1)} \tag{5}$$

$$= -\frac{\Delta}{\Delta+1} + \frac{e(\{u,w\}, P_{t-1})}{2(\Delta+1)} \tag{6}$$

where $e(A, B)$ is the number of edges between disjoint subsets $A, B \subseteq V$.

> **Remark:** *All our bounds study how c_t differs from c_{t-1}. So to be more precise, we should be conditioning on $\mathcal{E}_C \wedge \mathcal{E}_c$, where the latter is the event that c_{t-1} is some specific coloring c.*

We notice that $e(\{u,w\}, P_{t-1})$ can increase as more of the graph gets properly colored; but alongside, $(\mathbf{\Phi}_M + \mathbf{\Phi}_I)$ decreases. To account for this tradeoff between $(\mathbf{\Phi}_M + \mathbf{\Phi}_I)$ and $e(\{u,w\}, P_{t-1})$, we will include $\mathbf{\Phi}_P(t) = \Theta(e(V(I_t), P_t)/\Delta)$ in our potential function.

The expected behavior of $\mathbf{\Phi}_P$ at every recoloring step can also be bounded, but the calculations are somewhat tedious and can obscure the main idea. So we focus on the idea: when a vertex v has $|N(v) \cap (V(I_{t-1}) \cup P_{t-1})| = y$, then it can contribute $\leq (\Delta - y)$ new neighbors to $V(I_t) \cup P_t$. So overall, $N(v)$ appears in $\leq y(\Delta - y)$ new edges that contribute $f(y) := y(\Delta - y)/\Delta$ to $\mathbf{\Phi}_P(t) - \mathbf{\Phi}_P(t-1)$. This function is concave[2], so when $\mathbf{\Phi}_P$ becomes large, it begins to decrease fast.

[1] proof: induction on $|V(C)|$.

[2] visually, its plot looks like a dome.

Remark: *Formalizing the above uses a cool fact called Jensen's inequality, which can be stated in many different and equally interesting ways, depending on what kind of math one is doing.*

Since $\mathbf{\Phi}_P$ only decreases when $(\mathbf{\Phi}_M + \mathbf{\Phi}_I)$ becomes small, we are motivated to give a tiny constant to $\mathbf{\Phi}_P$. Our potential function, then, looks like:

$$\mathbf{\Phi}(t) = |M_t| + \frac{|I_t|}{10} + \frac{e(V(I_t), P_t)}{100\Delta}$$

Sadly, summing up the bounds on each term does not give a small enough δ for Theorem 1 to apply to $\mathbf{\Phi}$. It does, however, provide a *multiplicative* guarantee of decrease

$$\mathbf{Exp}[\mathbf{\Phi}(t) \mid \mathbf{\Phi}(t-1)] \;\le\; (1-\delta)\cdot\mathbf{Exp}[\mathbf{\Phi}(t-1)]$$

for a $\delta = \Theta(1/n)$. Happily, there a stronger drift theorem due to Doerr, Johannsen, and Winzen [DJW12] that works with multiplicative decrease, and gives us:

Theorem 2 (Multiplicative Drift of $\mathbf{\Phi}$)**.** $\mathcal{S}$ be all possible values that $\mathbf{\Phi}$ can take, and $s_{\min}$ be the smallest non-zero element in $\mathcal{S}$. Then, if $\exists\, \delta > 0$ such that $\forall s \in \mathcal{S}\setminus\{0\}$ and $\forall t > 0$, we had

$$\mathbf{Exp}[(\mathbf{\Phi}(t) - \mathbf{\Phi}(t-1)) \;\mid\; \mathbf{\Phi}(t-1) = s] \le -\delta s$$

then we can say

$$\mathbf{Exp}[\,T \;\mid\; \mathbf{\Phi}(0) = s_0] \le \frac{1 + \ln(s_0/s_{\min})}{\delta}$$

Intuitively, if $\mathbf{\Phi}(t)$ is multiplied by $(1-\delta)$ at each step, then it takes $\log_{1-\delta}(s_{\min}/s_0)$ steps for it to reach $s_{\min}$, whereafter its next decrease is to 0. We now must figure out $s_{\min}$ which, for a complicated function like $\mathbf{\Phi}$, can be a messy quantity. This motivates us to tweak[3] $\mathbf{\Phi}$ one last time.

- We know from Equation (3) that $\mathbf{\Phi}_M$ shows an expected additive decrease of $\approx 1/\Delta$. If at some t we get $\mathbf{\Phi}(t) < n/\Delta$, we also get $\mathbf{\Phi}_M(t) < n/\Delta$; then by Theorem 1, we expect only $O(n)$ more steps to be required. So once $\mathbf{\Phi}$ drops below n/Δ, it might as well be zero for our analysis. Hence, let

$$\mathbf{\Phi}'(t) = \begin{cases} \mathbf{\Phi}(t) & \text{if } \mathbf{\Phi}(t) \ge n/\Delta \\ 0 \text{ otherwise} & \end{cases}$$

$\mathbf{\Phi}'$ decreases at least as fast as $\mathbf{\Phi}$, so Theorem 2 applies to $\mathbf{\Phi}'$ as well. Then $s_{\min}$ becomes n/Δ, and since each term in $\mathbf{\Phi}(0)$ is $O(m)$, we get $s_0 = O(m) = O(n\Delta)$. So we have, using $\delta = \Theta(1/n)$,

$$\begin{aligned}\mathbf{Exp}[T \;\mid\; \mathbf{\Phi}'(0) = s_0] &= O\left(\frac{1 + \ln\left(\frac{n\Delta}{n/\Delta}\right)}{\delta}\right) \\ &= O(1 + 2\ln\Delta)\cdot\frac{1}{\Theta(1/n)} \\ &= O(n\log\Delta)\end{aligned}$$

as promised.

[3] "This isn't even my final form." – Frieza, in *Dragonball Z*.

Takeaway: *The Coupon Collector intuition helps us throughout. Indeed, that is why the* $(\Delta - y)$ *term was important – it arises from the* $(\Delta - \deg_C(v))$ *"unseen colors" around a vertex* v*. Moreover, the fact that we got* multiplicative *drift rather than additive, can also be explained via Coupon Collector – our progress towards seeing all coupons is slower when more coupons have been seen, and this slowdown brings in the* $\log$ *factor.*

It is also worth noting that $\mathbf{\Phi}$ *very clearly captures the changing configuration of the graph as the algorithm progresses:* $\mathbf{\Phi}_I$ *becomes important when* $\mathbf{\Phi}_M$ *becomes small, and* $\mathbf{\Phi}_P$ *becomes important when* $\mathbf{\Phi}_I$ *becomes small.*

Ponder This: *In our final reasoning, we used the worst possible* s_0*, ignoring the randomness of the initial coloring. In fact, the paper [BLM+20] goes on to assert that the initial coloring could have been* adversarial *without changing the result.*

References

[BLM+20] Daniel Bertschinger, Johannes Lengler, Anders Martinsson, Robert Meier, Angelika Steger, Miloš Trujić, and Emo Welzl. An optimal decentralized ($\delta + 1$)-coloring algorithm. In Fabrizio Grandoni, Grzegorz Herman, and Peter Sanders, editors, *28th Annual European Symposium on Algorithms (ESA 2020)*, volume 173 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 17:1–17:12, Dagstuhl, Germany, 2020. Schloss Dagstuhl–Leibniz-Zentrum für Informatik. 1, 3, 6

[DJW12] Benjamin Doerr, Daniel Johannsen, and Carola Winzen. Multiplicative drift analysis. *Algorithmica*, 64(4):673–697, feb 2012. 5

[HY04] Jun He and Xin Yao. A study of drift analysis for estimating computation time of evolutionary algorithms. *Natural Computing*, 3(1):21–35, 2004. 3