First-Fit Coloring of Forests in Random Arrival Model

Bartłomiej Bosek ✉ 🆔
Institute of Theoretical Computer Science, Faculty of Mathematics and Computer Science, Jagiellonian University, Kraków, Poland

Grzegorz Gutowski ✉ 🆔
Institute of Theoretical Computer Science, Faculty of Mathematics and Computer Science, Jagiellonian University, Kraków, Poland

Michał Lasoń ✉ 🆔
Institute of Mathematics of the Polish Academy of Sciences, ul. Śniadeckich 8, 00-656 Warszawa, Poland

Jakub Przybyło ✉ 🆔
AGH University of Krakow, Faculty of Applied Mathematics, al. A. Mickiewicza 30, 30-059 Kraków, Poland

──── **Abstract** ────

We consider a graph coloring algorithm that processes vertices in order taken uniformly at random and assigns colors to them using First-Fit strategy. We show that this algorithm uses, in expectation, at most $(1/2 + o(1)) \cdot \ln n / \ln \ln n$ different colors to color any forest with n vertices. We also construct a family of forests that shows that this bound is best possible.

1 Introduction

A *proper k-coloring* of a graph $G = (V, E)$ is a function $c : V \mapsto \{1, 2, \ldots, k\}$ that assigns a color $c(v)$ to each vertex $v \in V$ so that any adjacent vertices are colored differently, i.e., for each edge $\{u, w\} \in E$, $c(u) \neq c(w)$ is satisfied. For a given graph G, the smallest number k for which G admits a proper k-coloring is the *chromatic number* of G and is denoted by $\chi(G)$. Graph coloring is one of the most prominent disciplines within graph theory, with plenty of variants, applications, and deep connections to theoretical computer science. Coloring problems arise naturally in various job scheduling and resource allocation optimization scenarios.

The graph coloring problem is also very popular and well motivated in the online setting, with applications in job scheduling, dynamic storage allocation and resource management [9, 11, 12]. In the *online graph coloring* problem, an online algorithm receives as input a graph $G = (V, E)$ presented in an online fashion. The vertices of V are revealed one after one, in a *presentation order* $v_1 \ll v_2 \ll \ldots \ll v_n$. When a new vertex v_t is revealed in the t-th round, for $1 \leqslant t \leqslant n$, all the edges connecting v_t with vertices in $V_{t-1} = \{v_1, \ldots, v_{t-1}\}$ are also revealed. An online algorithm $\mathcal{A}$ has to immediately assign a feasible color to v_t, that is, a color that is different from those assigned to the neighbors of v_t in V_{t-1}. The goal is to minimize the total number of colors used.

The performance of an online graph coloring algorithm $\mathcal{A}$ is often measured by its *competitive ratio*. If we denote by $\chi_{\mathcal{A}}(G, \ll)$ the number of colors used by $\mathcal{A}$ when coloring a graph G in a presentation order $\ll$, then the competitive ratio of $\mathcal{A}$ is the maximum ratio $\chi_{\mathcal{A}}(G, \ll) / \chi(G)$ over all graphs G and all orders $\ll$. This means, that for a single graph, the analysis focuses on the worst case scenario, and the presentation order is often considered to be selected by an adversary. If $\mathcal{A}$ is a randomized algorithm, then it is usual to measure the competitive ratio with respect to the expected number of colors used when coloring G in order $\ll$ over all the random choices of $\mathcal{A}$.

Contrary to the offline setting, even coloring 2-colorable graphs is not easy for online algorithms. It is impossible to construct an online algorithm that would use any constant number of colors to color all 2-colorable graphs. Thus, the research focuses on restricted graph classes, and competitive ratio is often expressed as a function of the number of vertices in a graph. In fact, even for inputs restricted to online forests with n vertices, the best competitive ratio that can be achieved by either randomized or deterministic online coloring algorithm is $\Theta(\log n)$ colors [3, 6, 1]. The optimal number of colors used by any deterministic online algorithm to color 2-colorable graphs with n vertices is known to be somewhere between $2 \log_2 n - 10$ (by a result of Gutowski et al. [5]) and $2 \log_2 n$ (see a paper by Lovász, Saks, and Trotter [10]).

An algorithm known as First-Fit is arguably the simplest and the most understood of all online deterministic coloring algorithms. When a vertex v_t is revealed in the t-th round, First-Fit picks the least positive integer i which does not occur as a color of any of the previously colored neighbors of v_t and assigns i as a color of v_t. First-Fit performs well on online trees, achieving the competitive ratio of $\Theta(\log n)$ within this class. To be more precise, Bean [3] and, independently, Gyárfás and Lehel [6] proved that First-Fit uses at most $\log_2 n + 1$ colors on forests with n vertices. Irani [8] generalized this result, showing that First-Fit uses $O(d \log n)$ colors on d-degenerated graphs with n vertices. Later, Balogh et. al. [2] with corrections of Chang and Hsu [4] improved the above result to at most $(\log_{\frac{d+1}{d}} n + 2)$ colors. For bipartite graphs, there is an easy construction by Lovász, Saks, and Trotter [10] that shows that First-Fit uses as many as $\frac{n}{2}$ colors to color a 2-colorable graph with n vertices.

In this paper, we are interested in the *random arrival model* for online algorithms, that tries to focus on the average case, rather than the worst case scenario. In this model, the presentation order of a graph is not determined by an adversary, but instead it is selected uniformly at random from all possible permutations of the vertex set. The *performance ratio* of a deterministic online algorithm $\mathcal{A}$ on a graph G is measured by the expectation $\mathbb{E}_{\ll} \left[\frac{\chi_{\mathcal{A}}(G, \ll)}{\chi(G)} \right]$ over the random choice of $\ll$. The performance ratio of $\mathcal{A}$ on a class of graphs $\mathcal{G}$ is the the maximum performance ratio taken over all graphs in $\mathcal{G}$.

In particular, our intention is to start the systematic study of the performance ratio of First-Fit coloring algorithm in random arrival model. For a class of graphs $\mathcal{G}$, let $\mathrm{RFF}_{\mathcal{G}}(n)$ denote the maximum performance ratio of First-Fit taken over all graphs in $\mathcal{G}$ with n vertices, i.e.,

$$\mathrm{RFF}_{\mathcal{G}}(n) = \max_{G \in \mathcal{G}, |G|=n} \mathbb{E}_{\ll} \left[\frac{\chi_{\mathrm{FF}}(G, \ll)}{\chi(G)} \right].$$

For an easy comparison to the adversarial model, we define:

$$\mathrm{FF}_{\mathcal{G}}(n) = \max_{G \in \mathcal{G}, |G|=n} \max_{\ll} \left[\frac{\chi_{\mathrm{FF}}(G, \ll)}{\chi(G)} \right].$$

In this paper, we prove the following result, which establishes the performance of First-Fit on the class of forests.

▶ **Theorem 1.** *For the class $\mathcal{F}$ of forests, we have* $\mathrm{RFF}_{\mathcal{F}}(n) = (1/2 \pm o(1)) \cdot \ln n \, / \, \ln \ln n$.

As mentioned above, we have $\mathrm{FF}_{\mathcal{F}}(n) = \Theta(\log n)$. Theorem 1 shows that the randomization of the presentation order gives a noticeable, yet rather moderate increase in performance compared to the adversarial model.

One can also consider the following natural off-line graph coloring algorithm. Given a graph G, it selects an order $\ll$ of vertices of G uniformly at random. Then it uses First-Fit strategy to color vertices of G in order $\ll$. Theorem 1 shows that this algorithm uses $O(\log n \, / \, \log \log n)$ colors in expectation to color any forest. This algorithm does not compete with easy 2-coloring algorithms based on graph traversal. Our objective, however, is to establish the groundwork for a systematic analysis of First-Fit, and other simple online algorithms, in the random arrival model for other graph classes, such as d-degenerate graphs, bipartite graphs, and subsequently, k-colorable graphs. It should be noted that for 3-colorable graphs, the most effective known randomized online algorithm utilizes expected $O(n^{1/2})$ colors, as proved by Halldórsson [7]. Considering that the best lower bound for 3-colorable graphs is in the order of $\Omega(\log^2 n)$, as demonstrated by Vishwanathan [13], it appears that the direction we are pursuing holds promise for intriguing outcomes.

Another rationale for this line of inquiry is that First-Fit in random arrival model serves as an illustration for the following distributed coloring algorithm that works in the synchronous model. Each vertex of the graph represents a computation node, and edges represent communication links. To divide nodes into independent subsets, each node performs a sleep for a random number of units of time, and then assigns to itself the first possible number not assigned to any of the neighbors. This algorithm uses a small number of messages and is quite fast. Although we refrain from further delving into this setting, we underscore the versatility afforded by our analysis.

The paper is organized as follows. In Section 2 we show that First-Fit uses at most $(1/2 + o(1)) \cdot \ln n \, / \, \ln \ln n$ different colors in expectation on any forest with n vertices. In Section 3 we construct a family of trees for which First-Fit uses $(1/2 - o(1)) \cdot \ln n \, / \, \ln \ln n$ different colors in expectation. The final section contains brief comments and some open problems regarding the analyzed problem.

2 Upper bound

In this section we show that First-Fit uses at most $O(\log n \, / \, \log \log n)$ different colors in expectation on any forest with n vertices. For this purpose let us first make a basic observation. For any fixed graph $G = (V, E)$ and any order $\ll$ of V, we denote by $G^{\ll}$ the acyclic directed graph obtained from G by orienting every edge $\{u, v\} \in E$ so that it is oriented from u to v if and only if $u \ll v$.

▶ **Observation 2.** *For any graph $G = (V, E)$, a presentation order $\ll$ of V, a positive integer i, and a vertex $v \in V$, if First-Fit assigns color i to v when coloring G in order $\ll$ then there is a directed path in $G^{\ll}$ with i vertices and v as the last vertex.*

Proof. We prove the observation by induction with respect to i. The statement is trivial for $i = 1$, so let us assume $i \geqslant 1$ and that the thesis holds for all values up to i. Suppose that First-Fit assigns color $i + 1$ to vertex v in some round of the algorithm. At this point, v has a neighbor u with $u \ll v$ that was assigned with color i in some earlier round. By induction

hypothesis, there is a directed path P in $G^{\ll}$ with i vertices and u as the last vertex. Since $u \ll v$, and $G^{\ll}$ is acyclic we know that v does not belong to P. Thus, we get the desired path by extending P with directed edge (u, v).   ◄

We call a simple path P in a directed graph to be *bidirected* if it is either a directed path, or P can be split into two directed paths, each starting at one of the end points of P, and both sharing the same last vertex.

▶ **Corollary 3.** *For any forest $T = (V, E)$, a presentation order $\ll$ of V, and a positive integer $i \geqslant 2$, if First-Fit uses color i when coloring T in order $\ll$ then there is a bidirected path in $T^{\ll}$ with $2i - 2$ vertices.*

Proof. For $i = 2$, by Observation 2 we get a directed path with 2 vertices. Suppose $i \geqslant 3$, and that First-Fit assigns color i to some vertex v of T. At this point, v has a neighbor u that is colored $i - 1$, and a different neighbor w that is colored $i - 2$. By Observation 2 there is a directed path with $i - 1$ vertices and u as the last vertex, and a directed path with $i - 2$ vertices and w as the last vertex. As T is a forest, these paths are vertex disjoint. Since $u \ll v$, and $w \ll v$, we get the desired bidirected path in $T^{\ll}$ by extending both paths with directed edges (u, v) and (w, v).   ◄

▶ **Lemma 4.** *For the class $\mathcal{F}$ of forests, every $n \geqslant 3$, and an $\alpha_n = \frac{\ln \ln \ln n + 1}{\ln \ln n - \ln \ln \ln n - 1}$, we have:*

$$\mathrm{RFF}_{\mathcal{F}}(n) \leqslant \frac{(1 + \alpha_n) \ln n}{2 \ln \ln n} + \frac{3}{2}.$$

Proof. It is straightforward to verify that $\mathrm{RFF}_{\mathcal{F}}(3) = 1$ and $\mathrm{RFF}_{\mathcal{F}}(4) < 2$, hence the lemma holds for $n = 3, 4$. Let us assume that $n \geqslant 5$. Then, since the function $\ln \ln x - \ln \ln \ln x - 1 > 0$ for every $x > e$ except $x = e^e$, then $\ln \ln n - \ln \ln \ln n - 1 > 0$, and hence $\alpha_n > 0$. Consider any forest T with n vertices. Let $k = \left\lceil \frac{(1 + \alpha_n) \ln n}{\ln \ln n} \right\rceil$, and observe that our goal is to prove that First-Fit uses at most $k + 2$ colors in expectation when coloring T in a random order. Indeed, we can assume that $\chi(T) = 2$, as otherwise T is an independent set, and First-Fit uses only one color when coloring T in any order. Note that $(1 + \alpha_n) \ln n / \ln \ln n > \ln n / \ln \ln n > 1$. Hence, $k \geqslant 2$. By Corollary 3, for every order $\ll$ of the vertices of T and a positive integer $i \geqslant 2$, if $\chi_{FF}(T, \ll) = i$, then there is a bidirected path with $2i - 2$ vertices in $T^{\ll}$. Consider any two vertices x, y of T. There is at most one simple path in T with end points x and y. If this path exists, and is a path with exactly $2i - 2$ vertices, then the probability that this path is bidirected in $T^{\ll}$ for a random order $\ll$ is exactly $2^{2i-3} / (2i - 2)!$. Otherwise, the probability that there is such a path with end points x and y equals 0.

Thus, by the union bound, the probability that there is at least one bidirected path with $2i - 2$ vertices in $T^{\ll}$ is upper bounded by $\left(n^2 / 2\right) \cdot \left(2^{2i-3} / (2i - 2)!\right)$. Hence, the expected

number of colors used by First-Fit in a random order satisfies:

$$\mathbb{E}_{\ll}[\chi_{FF}(T, \ll)] = \sum_{i=1}^{\infty} i \cdot \mathbb{P}(\chi_{FF}(T, \ll) = i) \leqslant k + 1 + \sum_{i=k+2}^{\infty} \frac{i \cdot n^2 \cdot 2^{2i-4}}{(2i-2)!} \leqslant$$

$$= k + 1 + \frac{n^2 \cdot 4^{k-1}}{(2k)!} \cdot \sum_{i=k+2}^{\infty} \frac{i \cdot 2^{2(i-(k+1))}}{(2k+1) \cdot (2k+2) \cdot \ldots \cdot (2i-3) \cdot (2i-2)} \leqslant$$

$$\leqslant k + 1 + \frac{n^2 \cdot 4^{k-1}}{(2k)!} \cdot \sum_{i=k+2}^{\infty} \frac{2^{2(i-(k+1))}}{1 \cdot (2k+2) \cdot \ldots \cdot (2i-3) \cdot 2} \leqslant$$

$$\leqslant k + 1 + \frac{n^2 \cdot 4^{k-1}}{(2k)!} \cdot \sum_{i=1}^{\infty} \frac{2^{2i-1}}{(2i-1)!} =$$

$$= k + 1 + \frac{n^2 \cdot 4^{k-1}}{(2k)!} \cdot \sinh 2 \leqslant$$

$$\leqslant k + 1 + \frac{n^2 \cdot 4^{k}}{(2k)!}. \tag{1}$$

For $k = \left\lceil \frac{(1+\alpha_n)\ln n}{\ln\ln n} \right\rceil$, as $x \ln x - x$ is an increasing function for $x \geqslant 1$, and $\alpha_n > 0$, $k \geqslant 2$, we have that

$$\ln\left(\frac{(2k)!}{4^k}\right) \geqslant 2k\ln(2k) - 2k - k\ln 4 = 2k\ln k - 2k \geqslant$$

$$\geqslant 2 \cdot \frac{(1+\alpha_n)\ln n}{\ln\ln n} \cdot \ln\left(\frac{(1+\alpha_n)\ln n}{\ln\ln n}\right) - 2 \cdot \frac{(1+\alpha_n)\ln n}{\ln\ln n} \geqslant$$

$$\geqslant 2 \cdot \frac{(1+\alpha_n)\ln n}{\ln\ln n} \cdot \left(\ln\left(\frac{\ln n}{\ln\ln n}\right) - 1\right) =$$

$$= 2 \cdot \frac{\ln\ln n}{\ln\ln n - \ln\ln\ln n - 1} \cdot \frac{\ln n}{\ln\ln n} \cdot (\ln\ln n - \ln\ln\ln n - 1) =$$

$$= \ln(n^2),$$

and as a consequence, $\frac{(2k)!}{4^k} \geqslant n^2$, hence by (1),

$$\mathbb{E}_{\ll}[\chi_{FF}(T, \ll)] \leqslant k + 2.$$

As this holds for any forest T with n vertices, we conclude that

$$\mathrm{RFF}_{\mathcal{F}}(n) \leqslant \frac{(1+\alpha_n)\ln n}{2\ln\ln n} + \frac{3}{2},$$

which ends the proof. ◀

As $\alpha_n = O(\log\log\log n\ /\ \log\log n)$ in Lemma 4, we immediately get the following corollary.

▶ **Corollary 5.**

$\mathrm{RFF}_{\mathcal{F}}(n) \leqslant (1/2 + o(1)) \cdot \ln n\ /\ \ln\ln n$.

3   Lower bound

In this section, for any given $0 < \gamma < 1$, and a positive integer k we construct a tree T, such that First-Fit uses k colors to color T in a random presentation order with probability at least $1 - \gamma$. Thus, $\mathbb{E}_{\ll}[\chi_{FF}(T, \ll)] \geqslant k(1 - \gamma)$. The number of vertices in the constructed tree is of the order k^{dk} where d is a constant depending on γ. This implies that $\mathrm{RFF}_{\mathcal{F}}(n) = \Omega(\log n\ /\ \log\log n)$.

For the purpose of analysis, we use a slightly modified, yet equivalent random model. Namely, rather than choosing a permutation in a straightforward manner, we utilize a natural two-stage process. We first associate with every vertex v of a graph G an independent random variable $X_v \sim U[0, 1]$ uniformly distributed over $[0, 1]$ interval. We call X_v to be the *position* of a vertex v. Then, we order the vertices according to their positions, i.e., so that $u \ll v$ if and only if $X_u \leqslant X_v$. Note that such an order is uniquely determined with probability 1. Moreover, by the symmetry of the process, each resulting vertex permutation is equiprobable.

▶ **Lemma 6.** *For the class $\mathcal{F}$ of forests, we have:*

$\mathrm{RFF}_{\mathcal{F}}(n) \geqslant (1/2 - o(1)) \cdot \ln n\ /\ \ln\ln n$.

Proof. Fix any $0 < \gamma < 1$, and any integer $k \geqslant 3$. Let $c = {}^{10}/\gamma^2$ and $r = \lceil ck\ln k\rceil$. Set $\varepsilon_i = {}^{i\gamma}/k$ for $i = 1, 2, \ldots, k$. We recursively define rooted trees $T_1^r, T_2^r, \ldots, T_k^r$ as follows. The tree T_1^r is a single vertex, and for $i = 1, 2, \ldots, k-1$, the tree T_{i+1}^r is constructed of r copies of each of the trees T_j^r with $j = 1, 2\ldots, i$ by joining their roots to a single additional vertex – the root of T_{i+1}^r (hence the root of T_{i+1}^r has degree ri). See Figure 1.

Consider First-Fit coloring of T_i^r in a random presentation order, for any fixed $i \in \{1, 2, \ldots, k\}$. We assume that the presentation order is given by the positions $X_v \in [0, 1]$ drawn uniformly at random for every vertex v of T_i^r. By the construction of T_i^r, the longest simple path ending at the root of T_i^r includes at most i vertices. From Observation 2 we immediately obtain the following claim.

▷ **Claim 7.**   The color assigned by First-Fit to the root of T_i^r when coloring T_i^r in any order, does not exceed i.
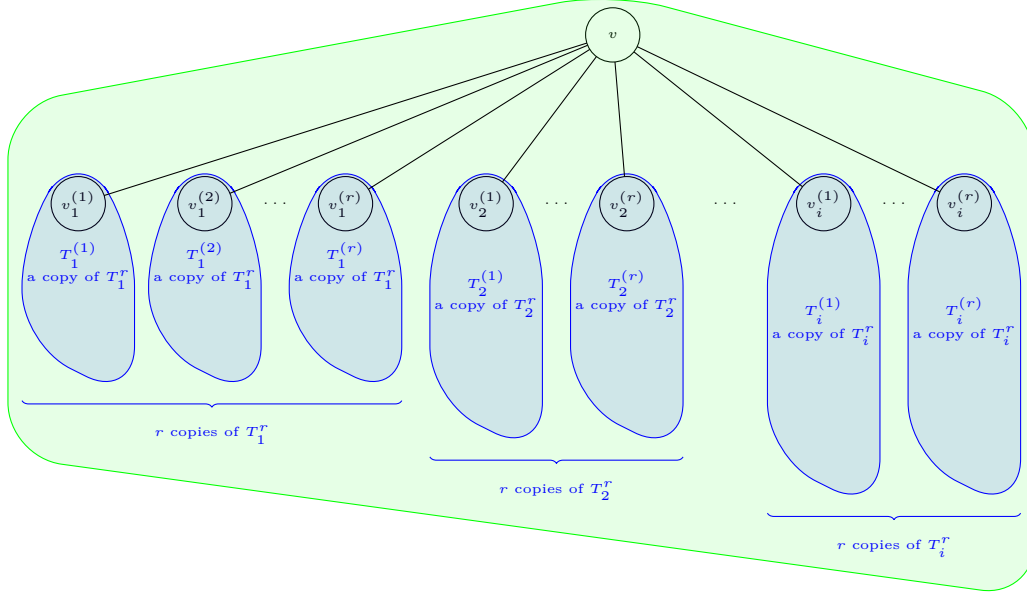
We define B_i to be the random event that First-Fit assigns color smaller then i to the root vertex of T_i^r when coloring T_i^r in a random presentation order.

▷ **Claim 8.**   For every $i = 1, 2, \ldots, k$ we have

$$\mathbb{P}(B_i) \leqslant \varepsilon_i = \frac{i\gamma}{k}.$$

Proof. We prove the claim by induction on i. For $i = 1$, First-Fit assigns color 1 to the only vertex of T_1^r and hence $\mathbb{P}(B_1) = 0$. Now, for the induction step, we fix any $1 \leqslant i \leqslant k-1$, and assume that

$$\mathbb{P}(B_j) \leqslant \varepsilon_j \tag{2}$$

Figure 1 Recursive construction of the tree T_{i+1}^r

holds for every $j = 1, 2, \ldots, i$. We shall prove that $\mathbb{P}(B_{i+1}) \leqslant \varepsilon_{i+1}$. For that we focus on the coloring of $T = T_{i+1}^r$. Denote the root of T by v and let $v_j^{(q)}$ with $j = 1, \ldots, i$, $q = 1, \ldots, r$ be the neighbors of v in T where each $v_j^{(q)}$ is the root of one of the r copies of T_j^r attached to v – we denote this copy as $T_j^{(q)}$. See Figure 1. We denote the (random) color First-Fit assigns to any vertex w by $c(w)$.

Suppose the position of v is fixed and equals x. Note that if for some $1 \leqslant j \leqslant i$ there is some q such that $c(v_j^{(q)}) \geqslant j$ and the position of $v_j^{(q)}$ is smaller than x, then $c(v) \neq j$. Indeed, due to Observation 2 and the fact that $v_j^{(q)}$ is positioned before v we have that $c(v_j^{(q)}) \leqslant j$ and it is assigned with color j by the assumption $c(v_j^{(q)}) \geqslant j$. Therefore, if there is such a q for every $1 \leqslant j \leqslant i$, then we get $c(v) \geqslant i + 1$. This allows for the following inequality,

$$\mathbb{P}\left(\overline{B_{i+1}} \mid X_v = x\right) \geqslant \mathbb{P}\left(\forall j \leqslant i \; \exists q \leqslant r : c(v_j^{(q)}) \geqslant j \wedge X_{v_j^{(q)}} < x \mid X_v = x\right) =$$

$$= \prod_{j=1}^{i} \mathbb{P}\left(\exists q \leqslant r : c(v_j^{(q)}) \geqslant j \wedge X_{v_j^{(q)}} < x \mid X_v = x\right) =$$

$$= \prod_{j=1}^{i} \left(1 - \prod_{q=1}^{r} \mathbb{P}\left(c(v_j^{(q)}) < j \vee X_{v_j^{(q)}} > x \mid X_v = x\right)\right) \tag{3}$$

where the two last equalities above follow by the independence of the corresponding events for a fixed value of x. Obviously $\mathbb{P}(X_{v_j^{(q)}} > x \mid X_v = x) = 1 - x$, as positions of the vertices are independent. Further, for any assignment X of positions to all vertices of T, one can consider First-Fit coloring of $T_j^{(q)}$ in order given by the restriction of X to the vertices of $T_j^{(q)}$. Color assigned to the root of $T_j^{(q)}$ in this restricted coloring is not greater then the

color assigned to this vertex in the coloring of T. Thus, $\mathbb{P}(c(v_j^{(q)}) < j \mid X_v = x) \leqslant \mathbb{P}(B_j)$, and by (2), for every $j \leqslant i$, we have:

$$\mathbb{P}\left(c(v_j^{(q)}) < j \vee X_{v_j^{(q)}} > x \mid X_v = x\right) \leqslant \max\left\{0, 1 - x + \varepsilon_j\right\}. \tag{4}$$

By (3) and (4), we thus obtain that

$$\mathbb{P}(B_{i+1}) = 1 - \mathbb{P}\left(\overline{B_{i+1}}\right) = 1 - \int_0^1 \mathbb{P}\left(\overline{B_{i+1}} \mid X_v = x\right) dx \leqslant$$

$$\leqslant 1 - \int_0^1 \prod_{j=1}^i \max\left\{0, 1 - (1 - x + \varepsilon_j)^r\right\}\} dx \leqslant$$

$$\leqslant 1 - \int_{\varepsilon_i}^1 (1 - (1 - x + \varepsilon_i)^r)^i dx, \tag{5}$$

where the last inequality follows by the fact that $\varepsilon_j \leqslant \varepsilon_i$ for $j \leqslant i$. Substituting $y = 1 - x + \varepsilon_i$ in (5) we further obtain:

$$\mathbb{P}(B_{i+1}) \leqslant 1 - \int_{\varepsilon_i}^1 (1 - y^r)^i dy. \tag{6}$$

Let $f(y) = (1 - y^r)^i$. Observe that $f(0) = 1$, $f(1) = 0$, and f is strictly decreasing in $[0, 1]$. Thus, we obtain that $\int_0^{\varepsilon_i} f(y) dy \leqslant \varepsilon_i$, and by (6),

$$P(B_{i+1}) \leqslant \varepsilon_i + 1 - \int_0^1 f(y) dy \leqslant \varepsilon_i + 1 - \left(1 - \frac{\gamma}{2k}\right) \cdot f\left(1 - \frac{\gamma}{2k}\right) \leqslant$$

$$\leqslant \varepsilon_i + \frac{\gamma}{2k} + 1 - f\left(1 - \frac{\gamma}{2k}\right). \tag{7}$$

In order to prove that $P(B_{i+1}) \leqslant \varepsilon_{i+1} = \varepsilon_i + \frac{\gamma}{k}$ it thus remains to show that $f(1 - \frac{\gamma}{2k}) \geqslant 1 - \frac{\gamma}{2k}$. Note that

$$f\left(1 - \frac{\gamma}{2k}\right) = \left(1 - \left(1 - \frac{\gamma}{2k}\right)^r\right)^i \geqslant \left(1 - \left(1 - \frac{\gamma}{2k}\right)^{ck \ln k}\right)^k = \left(1 - \left(1 - \frac{\gamma}{2k}\right)^{\frac{2k}{\gamma} \frac{\gamma}{2} c \ln k}\right)^k.$$

Now, for $\alpha = \frac{2k}{\gamma} \geqslant 1$ we have that $(1 - \frac{1}{\alpha})^\alpha < \frac{1}{e}$. Thus,

$$f\left(1 - \frac{\gamma}{2k}\right) \geqslant \left(1 - \frac{1}{e^{\frac{\gamma c \ln k}{2}}}\right)^k = \left(1 - \frac{1}{k^{\frac{\gamma c}{2}}}\right)^k.$$

As for $0 < \beta = \frac{1}{k^{\frac{\gamma c}{2}}} < 1$ we have that $1 - \beta > e^{-\frac{\beta}{1-\beta}}$, we thus further obtain that

$$f\left(1 - \frac{\gamma}{2k}\right) \geqslant e^{-\frac{\beta k}{1-\beta}} = e^{-\frac{k}{k^{\frac{\gamma c}{2}} - 1}} \geqslant 1 - \frac{k}{k^{\frac{\gamma c}{2}} - 1}.$$

Now, for $c = \frac{10}{\gamma^2}$, and using $k \geqslant 3$ we get

$$f\left(1 - \frac{\gamma}{2k}\right) \geqslant 1 - \frac{k}{k^{\frac{5}{\gamma}} - 1} \geqslant 1 - \frac{1}{k^{\frac{3}{\gamma}}}.$$

Observe that $k^{\frac{3}{\gamma}} > \frac{2k}{\gamma}$ for $\gamma \in [0, 1]$, and $k \geqslant 3$. This finally gives that $f(1 - \frac{\gamma}{2k}) \geqslant 1 - \frac{\gamma}{2k}$, and consequently, by (7), $P(B_{i+1}) \leqslant \varepsilon_{i+1}$, which ends the proof of Claim 8.   ◀

By Claims 7 and 8, First-Fit assigns color k to the root of T_k^r with probability at least $1 - \varepsilon_k$. Using Claim 7 (or Observation 2), one can thus easily deduce the following claim.

$\triangleright$ **Claim 9.** $\mathbb{P}(\chi_{FF}(T_k^r, \ll) = k) \geqslant 1 - \gamma$.

$\triangleright$ **Claim 10.** T_k^r has exactly $(r + 1)^{k-1}$ vertices.

Proof. We prove the claim by induction on k. For $k = 1$ the claim trivially holds. Let us assume that $k \geqslant 1$ and that the claim holds for all the trees $T_1^r, \ldots, T_k^r$. By the construction of T_{k+1}^r,

$$|T_{k+1}^r| = 1 + r \cdot \sum_{i=1}^k |T_i^r| = 1 + r \cdot \sum_{i=1}^k (r+1)^{i-1} = 1 + r \cdot \frac{1 - (r+1)^k}{1 - (r+1)} = (r+1)^k.$$

◀

For any value of k, we set $\gamma = \frac{1}{\ln k}$, and for this choice of γ we get $c = 10 \ln^2 k$ and $r = \lceil 10k \ln^3 k \rceil$. Let us denote by n_k the number of vertices in T_k^r. Then the following hold for k large enough. Firstly, by Claim 10, we have $\ln n_k \geqslant k$. Consequently, again by Claim 10,

$$\ln n_k \leqslant (k-1) \ln \left(2 + 10k \ln^3 k\right) \leqslant k \left(\ln k + 4 \ln \ln k\right)$$

$$\leqslant k \left(\ln \ln n_k + 4 \ln \ln \ln n_k\right). \tag{8}$$

By Claim 9 and (8), we thus finally obtain that

$$\mathbb{E}_{\ll}\left[\chi_{FF}(T_k^r, \ll)\right] \geqslant k \cdot \left(1 - \frac{1}{\ln k}\right) \tag{9}$$

$$\geqslant \frac{\ln n}{\ln \ln n + 4 \ln \ln \ln n} \cdot \left(1 - \frac{1}{\ln \ln n - 2 \ln \ln \ln n}\right) = g(n). \tag{10}$$

Note that $g(n)$ is an increasing function for large enough n. Note also that n_k is also an increasing function of k. Consider any (large enough) n such that $n_k \leqslant n \leqslant n_{k+1}$ for some k. Then, since $\mathrm{RFF}_{\mathcal{F}}(n)$ is a nondecreasing function of n, by (9) and (10),

$$\mathrm{RFF}_{\mathcal{F}}(n) \geqslant \mathrm{RFF}_{\mathcal{F}}(n_1) \geqslant \frac{1}{2} \cdot \mathbb{E}_{\ll}\left[\chi_{FF}(T_k^r, \ll)\right] \geqslant \frac{1}{2} \cdot k \cdot \left(1 - \frac{1}{\ln k}\right)$$

$$\geqslant \frac{1}{2} \cdot \left((k+1) \cdot \left(1 - \frac{1}{\ln(k+1)}\right) - 1\right) \geqslant \frac{1}{2} \cdot (g(n_{k+1}) - 1)$$

$$\geqslant \frac{1}{2} \cdot (g(n) - 1) = \left(\frac{1}{2} - o(1)\right) \frac{\ln n}{\ln \ln n},$$

which finishes the proof of Lemma 6. ◀

4   Final comments

Finally, Theorem 1 is obtained by combining the matching bounds of Corollary 5 and Lemma 6. This concludes our investigation of the efficiency of First-Fit in the random arrival model on the class of forests. First-Fit is efficient on this class even in the adversarial model. Still, the randomization of the presentation order allows for some increase in performance. This raises the hope that First-Fit is more effective in the average case than it is in the worst case also on some other graph classes.

The systematic analysis of First-Fit in the random arrival model on other graph classes should continue for the class $\mathcal{B}$ of bipartite graphs. There, First-Fit is known to be extremely inefficient in the adversarial model with $\mathrm{FF}_{\mathcal{B}} = \Theta(n)$. However, there is another simple algorithm [10], a clever modification of First-Fit, that uses $O(\log n)$ colors to color any bipartite graph with n vertices. A natural extension of our research would be to assess the First-Fit algorithm in the random arrival model on bipartite graphs. We anticipate that the outcome will not deviate significantly from the results obtained for forests.

▶ **Conjecture 11.** *First-Fit in the random arrival model uses at most* $\mathrm{poly}(\log n)$ *colors when coloring any bipartite graph with n vertices.*

Addressing the aforementioned conjecture represents a significant cognitive pursuit. However, a truly remarkable feat would be to establish some nontrivial bounds for 3-colorable graphs. We want to thank Anna Zych-Pawlewicz for inspiring the work on this project.

References

1   Susanne Albers and Sebastian Schraink. Tight bounds for online coloring of basic graph classes. In *ESA 2017: 25th Annual European Symposium on Algorithms, Vienna, Austria, September 4-8, 2017, Proceedings*, volume 87 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 7:1–7:14. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2017. `doi: 10.4230/LIPIcs.ESA.2017.7`.

2   József Balogh, Stephen G. Hartke, Qi Liu, and Gexin Yu. On the first-fit chromatic number of graphs. *SIAM Journal on Discrete Mathematics*, 22(3):887–900, 2008. `doi:10.1137/ 060672479`.

3   Dwight R. Bean. Effective coloration. *The Journal of Symbolic Logic*, 41(2):469–480, 1976. `doi:10.2307/2272247`.

4   Gerard Jennhwa Chang and Hsiang-Chun Hsu. First-fit chromatic numbers of *d*-degenerate graphs. *Discrete Mathematics*, 312(12-13):2088–2090, 2012. `doi:10.1016/j.disc.2012.03. 029`.

5   Grzegorz Gutowski, Jakub Kozik, Piotr Micek, and Xuding Zhu. Lower bounds for online graph colorings. In *ISAAC 2014: 25th International Symposium on Algorithms and Computation, Jeonju, Korea, December 15-17, 2014, Proceedings*, volume 8889 of *Lecture Notes in Computer Science*, pages 507–515. Springer, 2014. `doi:10.1007/978-3-319-13075-0_40`.

6   András Gyárfás and Jenő Lehel. On-line and first fit colorings of graphs. *Journal of Graph Theory*, 12(2):217–227, 1988. `doi:10.1002/jgt.3190120212`.

7   Magnús M. Halldórsson. Parallel and on-line graph coloring. *Journal of Algorithms*, 23(2):265–280, 1997. `https://doi.org/10.1006/jagm.1996.0836`. `doi:10.1006/jagm.1996.0836`.

8   Sandy Irani. Coloring inductive graphs on-line. *Algorithmica*, 11(1):53–72, 1994. `doi: 10.1007/BF01294263`.

9   Hal A. Kierstead. Coloring graphs on-line. In *Online Algorithms: The State of the Art*, pages 281–305. Springer Berlin Heidelberg, Berlin, Heidelberg, 1998. `doi:10.1007/BFb0029574`.

10   László Lovász, Michael Saks, and W. T. Trotter. An on-line graph coloring algorithm with sublinear performance ratio. *Discrete Mathematics*, 75(1-3):319–325, 1989. Graph theory and combinatorics (Cambridge, 1988). `doi:10.1016/0012-365X(89)90096-4`.

11   Dániel Marx. Graph colouring problems and their applications in scheduling. *Periodica Polytechnica Electrical Engineering*, 48(1-2):11–16, 2004. URL: `https://pp.bme.hu/ee/article/view/926`.

12   Lata Narayanan. Channel assignment and graph multicoloring. In *Handbook of Wireless Networks and Mobile Computing*, chapter 4, page 71–94. John Wiley & Sons, Inc., USA, 2002. `doi:10.1002/0471224561.ch4`.

13   Sundar Vishwanathan. Randomized online graph coloring. *Journal of Algorithms. Cognition, Informatics and Logic*, 13(4):657–669, 1992. `https://doi.org/10.1016/0196-6774(92)90061-G`. `doi:10.1016/0196-6774(92)90061-G`.