

Reinforcement learning for graph theory,

II. Small Ramsey numbers*

Mohammad Ghebleh[†] , Salem Al-Yakoob , Ali Kanso 

Department of Mathematics, Faculty of Science, Kuwait University, Safat, Kuwait

Dragan Stevanović[‡] 

College of Integrative Studies, Abdullah Al Salem University, Khaldiya, Kuwait

Received 29 March 2024, accepted 29 December 2024, published online 25 February 2025

Abstract

We describe here how the recent Wagner’s approach for applying reinforcement learning to construct examples in graph theory can be used in the search for critical graphs for small Ramsey numbers. We illustrate this application by providing lower bounds for the small Ramsey numbers $R(K_{2,5}, K_{3,5})$, $R(B_3, B_6)$ and $R(B_4, B_5)$ and by improving the lower known bound for $R(W_5, W_7)$.

Keywords: Ramsey number, critical graph, reinforcement learning, cross-entropy method.

Math. Subj. Class.: 05C55, 05D10, 68T20

1 Introduction

Due to discrete and easily programmable nature of graphs, computers were used, besides running usual graph algorithms, to also unearth peculiar examples of graphs answering various research questions for more than five decades. Several pieces of software were developed specifically to find extremal graphs for arbitrary invariant expressions or to conjecture relations between graph invariants. The most well-known among them are probably AutoGraphiX [1–4] and Graffiti [6–10], whose Dalmatian heuristic that selectively extends the starting sets of graphs and invariant expressions was recently reimplemented in [12].

More recently, Adam Wagner [15] has shown how the cross-entropy method, a particular reinforcement learning technique developed by Rubinstein [5, 14], can be applied to

*This work was supported and funded by Kuwait University Research Grant No. SM05/22.

[†]Corresponding author.

[‡]On leave from Mathematical Institute of Serbian Academy of Sciences and Arts.

E-mail addresses: mohammad.ghebleh@ku.edu.kw (Mohammad Ghebleh), salem.alyakoob@ku.edu.kw (Salem Al-Yakoob), ali.kanso@ku.edu.kw (Ali Kanso), dragan.stevanovic@asu.edu.kw (Dragan Stevanović)

construct graphs that maximize a pre-defined reward, which he used to find several counterexamples to conjectures in graph theory and combinatorics. We have reimplemented his approach in [11] in order to make it faster and more stable, and applied it to disprove a larger set of previously conjectured upper bounds on the Laplacian spectral radius of graphs.

For graphs $G_1, \dots, G_m$, the m -color Ramsey number $R(G_1, \dots, G_m)$ is defined as the least integer n such that for any edge-coloring of the complete graph K_n with m colors there exists some $1 \leq i \leq m$ such that the subgraph of K_n formed by the edges of color i contains a copy of G_i (not necessarily induced). An m -edge-coloring of K_n is critical if it does not contain a copy of G_i having edges of color i for any $i = 1, \dots, m$. The existence of a critical m -edge-coloring of K_n implies that $R(G_1, \dots, G_m) \geq n + 1$. See Radziszowski's dynamic survey [13] for a wealth of up-to-date information on the values and bounds for specific Ramsey numbers.

Here we show that, with a small modification, the cross-entropy method can also be used to find new examples of critical edge-colorings, using which we improve four lower bounds for small Ramsey numbers.

2 Cross-entropy method for small Ramsey numbers

The cross-entropy method is an optimization procedure that employs a feed-forward neural network to iteratively construct a new batch of m -edge-colorings of K_n (for fixed m and n) and to learn the structure of the best performing ones among them.

Edge-colorings are constructed via a sequence of observations and actions. Each observation is a partial m -edge-coloring of K_n together with a one-hot encoding of the current edge that needs to be colored, while the action is the color selected for the current edge. The neural network is set to accept an observation as an input for which, using a few hidden layers, it then outputs the probability distribution of m colors. The action (i.e., the color of the current edge) is chosen randomly according to this probability distribution, and the neural network is then used again with an updated observation to produce the probability distribution of colors for the next edge, until the ongoing m -edge-coloring of K_n is completed.

This process of constructing m -edge-colorings of K_n is repeated for a specified number of times, in order to obtain a full batch of edge-colorings. Each edge-coloring from the batch receives a reward that is equal to the sum of the numbers of copies of the graph G_i with edges of color i that it contains for $i = 1, \dots, m$. Obviously, critical edge-colorings are those that have a zero reward. At this point, a specified percentage of edge-colorings with smallest rewards are selected from the batch, and the neural network is trained on the sequence of observations and actions that led to these best edge-colorings. A smaller percentage of the very best edge-colorings is added to the next batch, to ensure that the best performing edge-colorings are not lost between two batches.

Training of the neural network on the best performing edge-colorings from each batch ensures that the edge-colorings in the next batch will tend to use actions that have previously led to the best performers, which will hopefully decrease the smallest reward in the next batch, due to the randomized process of selecting edge colors (the neural network determines the probability distribution of edge colors only). In any case, keeping a few best survivors from one batch to the next ensures that the smallest reward cannot increase.

Over time, the network will start to be retrained on repeated best performers and it will eventually almost surely lead to the same edge-colorings over and over again in the later batches, thus getting stuck at a local minimum. To postpone this inevitability, we have introduced a varying percentage of fully randomized actions, which increases whenever the smallest reward has not decreased for a longer time. This can help the neural network to avoid learning too specific parts of the structure too soon, hopefully leading it to find a better local minimum in the long term. If this local minimum happens to also be a global minimum with a zero reward, then we have arrived at a critical edge-coloring. Otherwise, the whole process simply has to be repeated from scratch, as the weights of the initial neural network are selected entirely randomly, so that the start of the new training session will very probably lead to a different local minimum.

A Python implementation of this method, with subgraph counting implemented in Java for speedup, is available at github.com/dragancel06/cema-ramsey. It is heavily commented so that the interested user will easily find her way through the code. Using it, we have been able to find four new critical edge-colorings, which are presented in the next section.

3 New lower bounds for Ramsey numbers

For two graphs G and H , the join $G + H$ is the graph obtained from the union of G and H by adding all the edges between a vertex of G and a vertex of H . The complete bipartite graph $K_{m,n}$ is the join $\overline{K_m} + \overline{K_n}$ of two empty graphs on m and n vertices, respectively. The wheel W_n is the join $K_1 + C_{n-1}$ of a central vertex and a cycle on $n - 1$ vertices. The book B_n is the join $K_2 + \overline{K_n}$, which has $n + 2$ vertices and can be visualised as n triangular pages having a common edge.

In the following theorems, we improve the known lower bound $R(W_5, W_7) \geq 13$ [13] and give the first lower bounds on $R(K_{2,5}, K_{3,5})$, $R(B_3, B_6)$ and $R(B_4, B_5)$ by identifying critical edge-colorings using the above cross-entropy method.

Theorem 3.1. $R(W_5, W_7) \geq 14$.

Proof. The following matrix gives a critical edge-coloring of K_{13} with colors 0 and 1 such that there is no W_5 with edges of color 0 and no W_7 with edges of color 1:

$$\begin{bmatrix} & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 & & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & \end{bmatrix}$$

Alternatively, the graph determined by this adjacency matrix (edge color 1) does not contain W_7 as a subgraph, while its complement (edge color 0) does not contain W_5 as a subgraph. $\square$

Theorem 3.2. $R(K_{2,5}, K_{3,5}) \geq 20$.

Proof. The following matrix gives a critical edge-coloring of K_{19} with colors 0 and 1 such that there is no $K_{2,5}$ with edges of color 0 and no $K_{3,5}$ with edges of color 1:

	0	1	1	1	1	0	0	1	1	1	0	1	1	0	0	1	0	0
0		1	0	0	1	1	1	0	1	1	1	0	1	1	1	0	0	1
1	1		0	0	1	1	1	1	1	0	1	0	1	0	0	0	1	0
1	0	0		1	1	0	0	1	1	1	1	1	0	1	1	0	1	0
1	0	0	1		0	1	0	0	1	0	1	1	1	1	1	0	0	1
1	1	1	1	0		0	1	1	0	1	1	0	0	1	0	0	0	1
0	1	1	0	1	0		0	1	1	0	1	1	1	0	1	0	0	1
0	1	1	0	0	1	0		1	0	0	0	1	1	1	0	1	1	1
1	0	1	1	0	1	1	1		0	0	0	1	0	0	1	1	0	1
1	1	1	1	1	0	1	0	0		0	0	0	1	0	1	1	1	1
1	1	0	1	0	1	0	0	0	0		1	1	1	1	1	1	0	0
0	1	1	1	1	1	1	0	0	0	1		0	1	1	0	0	1	0
1	0	0	1	1	0	1	1	1	0	1	0		1	1	1	0	0	0
1	1	1	0	1	0	1	1	0	1	1	1	1		0	0	1	0	0
0	1	0	1	1	1	0	1	0	0	1	1	1	0		1	0	1	1
0	1	0	1	0	0	1	0	1	1	1	0	1	0	1		1	1	1
1	0	0	0	0	0	0	1	1	1	1	0	0	1	0	1		1	1
0	0	1	1	1	0	0	1	0	1	0	1	0	0	1	1	1		0
0	1	0	0	1	1	1	1	1	1	0	0	0	0	1	1	1	0	

□

Theorem 3.3. $R(B_3, B_6) \geq 17$.

Proof. The following matrix gives a critical edge-coloring of K_{16} with colors 0 and 1 such that there is no B_3 with edges of color 0 and no B_6 with edges of color 1:

	1	1	1	1	0	1	1	0	0	1	1	0	1	0	0
1		0	0	1	1	1	1	1	1	0	1	1	0	0	0
1	0		1	1	0	1	0	0	0	0	1	1	1	0	1
1	0	1		0	0	0	1	1	0	1	1	1	1	1	0
1	1	1	0		0	1	0	1	1	1	0	1	0	0	1
0	1	0	0	0		1	1	1	1	0	1	0	1	1	1
1	1	1	0	1	1		1	1	0	1	0	0	0	1	0
1	1	0	1	0	1	1		0	1	1	0	1	0	1	1
0	1	0	1	1	1	1	0		0	1	1	1	0	1	1
0	1	0	0	1	1	0	1	0		1	1	1	1	0	1
1	0	0	1	1	0	1	1	1	1		0	0	1	1	0
1	1	1	1	0	1	0	0	1	1	0		1	1	0	0
0	1	1	1	1	0	0	1	1	1	0	1		0	0	1
1	0	1	1	0	1	0	0	0	1	1	1	0		1	1
0	0	0	1	0	1	1	1	1	0	1	0	0	1		1
0	0	1	0	1	1	0	1	1	1	0	0	1	1	1	

□

Theorem 3.4. $R(B_4, B_5) \geq 18$.

Proof. The following matrix gives a critical edge-coloring of K_{17} with colors 0 and 1 such that there is no B_4 with edges of color 0 and no B_5 with edges of color 1:

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}$$

□

4 Concluding remarks

Numerous experiments with the implemented method showed that higher learning percentages and larger batch sizes tend to help it to more easily identify promising directions for learning. While the results from the previous section do prove that the cross-entropy method may lead to new examples of critical edge-colorings for small Ramsey numbers, most often it actually leaves the questions simply unanswered.

As a prominent example, the best 2-edge-colorings of K_{17} that it managed to produce after hundreds of runs, have between 1 and 8 monochromatic copies of B_3 and B_6 , depending on the run. Regardless of the number of runs, one still cannot be certain whether the reason for this is because truly $R(B_3, B_6) = 17$ or we just did not attempt that one additional lucky run of the cross-entropy method?

Another question is how to predict when is the time to end the current run? We have observed cases when the reward reaches a plateau after 300–400 batches and then it remains constant until a sudden decrease more than 10,000 batches later. This decrease was apparently a consequence of a lucky randomization of actions that accidentally led to a smaller reward, moving away from the local minimum where it was stuck for thousands of batches. However, what this learning method really needs is an improved way of getting out of local minima. At present, a jump to a better local minimum through randomization of actions has to be made within a single batch, as otherwise such randomized colorings will not qualify within the surviving or even the learning percentage, and they will be already discarded by the time the next batch is constructed.

The final remark is that, although the cross-entropy method does manage to construct some critical edge-colorings, it may not actually be the best way for doing it. Namely, each

critical edge-coloring of K_n automatically yields a critical edge-coloring of K_{n-1} simply by deleting an arbitrary vertex of K_n . Thus, it may actually be more beneficial to proceed inductively with learning: assuming the machine learning model M_{n-1} knows how to construct critical edge-colorings of K_{n-1} to then somehow use the acquired knowledge of M_{n-1} to train the model M_n that will also construct critical edge-colorings of K_n , provided $n < R(G_1, \dots, G_m)$.

We leave the resolution of these questions for future work.

ORCID iDs

Mohammad Ghebleh  <https://orcid.org/0000-0003-2291-0892>

Salem Al-Yakoob  <https://orcid.org/0000-0001-6885-3198>

Ali Kanso  <https://orcid.org/0000-0002-4366-841X>

Dragan Stevanović  <https://orcid.org/0000-0003-2908-305X>

References

- [1] M. Aouchiche, J. M. Bonnefoy, A. Fidahoussen, G. Caporossi, P. Hansen, L. Hiesse, J. Lacheré and A. Monhait, Variable neighborhood search for extremal graphs. XIV: The AutoGraphiX 2 system, in: *Global Optimization. From Theory to Implementation.*, Springer, New York, NY, pp. 281–310, 2006.
- [2] M. Aouchiche, G. Caporossi, P. Hansen and M. Laffay, Autographix: a survey, in: *7th International Colloquium on Graph Theory, Hyeres, France, September 12–16, 2005*, Elsevier, Amsterdam, pp. 515–520, 2005, doi:10.1016/j.endm.2005.06.090, <https://doi.org/10.1016/j.endm.2005.06.090>.
- [3] M. Aouchiche, G. Caporossi, P. Hansen, N. Mladenović and C. Lucas, Finding conjectures in graph theory with AutoGraphiX, *Le Cahiers du GERAD* (2016), G–2016–90.
- [4] G. Caporossi and P. Hansen, Variable neighborhood search for extremal graphs. I: The AutoGraphiX system, *Discrete Math.* **212** (2000), 29–44, doi:10.1016/S0012-365X(99)00206-X, [https://doi.org/10.1016/S0012-365X\(99\)00206-X](https://doi.org/10.1016/S0012-365X(99)00206-X).
- [5] P.-T. De Boer, D. P. Kroese, S. Mannor and R. Y. Rubinstein, A tutorial on the cross-entropy method, *Ann. Oper. Res.* **134** (2005), 19–67, doi:10.1007/s10479-005-5724-z, <https://doi.org/10.1007/s10479-005-5724-z>.
- [6] S. Fajtlowicz, On conjectures of graffiti. II, in: *Combinatorics, Graph Theory, and Computing, Proc. 18th Southeast. Conf., Boca Raton, 1987*, Congr. Numerantium 60 (1987), pp. 189–197.
- [7] S. Fajtlowicz, On conjectures of graffiti. III, in: *Combinatorics, Graph Theory, and Computing, Proc. 19th Southeast. Conf., Boca Raton, 1988*, Congr. Numerantium 66 (1988), pp. 23–32.
- [8] S. Fajtlowicz, On conjectures of graffiti. IV, in: *Combinatorics, Graph Theory, and Computing, Proc. 20th Southeast Conf., Boca Raton, 1989*, Congr. Numerantium 70 (1990), pp. 231–240.
- [9] S. Fajtlowicz, On conjectures of Graffiti, *Discrete Math.* **72** (1988), 113–118, doi:10.1016/0012-365X(88)90199-9, [https://doi.org/10.1016/0012-365X\(88\)90199-9](https://doi.org/10.1016/0012-365X(88)90199-9).
- [10] S. Fajtlowicz, On conjectures of Graffiti. V, in: *Graph theory, Combinatorics, Algorithms and Applications, Kalamazoo, MI*, Wiley, New York, NY, pp. 367–376, 1995.
- [11] M. Ghebleh, S. Al-Yakoob, A. Kanso and D. Stevanović, Reinforcement learning for graph theory, I. Reimplementation of Wagner’s approach, 2024, [arXiv:2403.18429](https://arxiv.org/abs/2403.18429) [math.CO].
- [12] C. E. Larson and N. Van Cleemput, Automated conjecturing. I: Fajtlowicz’s Dalmatian heuristic revisited, *Artif. Intell.* **231** (2016), 17–38, doi:10.1016/j.artint.2015.10.002, <https://doi.org/10.1016/j.artint.2015.10.002>.

- [13] S. P. Radziszowski, Small Ramsey numbers, *Electron. J. Comb.* **DS01** (2021), doi:10.37236/21, <https://doi.org/10.37236/21>.
- [14] R. Y. Rubinstein, Optimization of computer simulation models with rare events, *Eur. J. Oper. Res.* **99** (1997), 89–112, doi:10.1016/S0377-2217(96)00385-2, [https://doi.org/10.1016/S0377-2217\(96\)00385-2](https://doi.org/10.1016/S0377-2217(96)00385-2).
- [15] A. Wagner, Constructions in combinatorics via neural networks, 2021, [arXiv:2104.14516](https://arxiv.org/abs/2104.14516) [math.CO].