

The Logic Behind Colour Refinement

Sandra Kiefer  

University of Oxford, UK

Abstract

Colour Refinement is a combinatorial algorithm that computes a vertex colouring for an input graph to reveal its structural properties. Each iteration of the algorithm refines the current colouring by assessing local information. More precisely, the new colour of a vertex is determined by its current colour and the multiset of colours in its neighbourhood. This refinement procedure continues until it reaches a stable partition of the vertex set into colour classes.

On the practical side, the algorithm admits fast implementations. Because the final colouring is isomorphism-invariant, executing the algorithm on two graphs in parallel can be used to demonstrate that they are not isomorphic. From a theoretical perspective, the algorithm is arguably the most straightforward combinatorial approach to detecting asymmetries—specifically for distinguishing vertices that do not belong to the same orbit of the automorphism group of the graph. Its numerous connections to other areas in computer science stand as evidence of its robustness and naturalness and make it a fascinating object of research.

Among the most elegant connections is the link to counting logic. Colour Refinement assigns distinct final colours to two vertices in a graph if and only if there is a formula in the two-variable fragment C^2 of the logic C that distinguishes them, meaning that the formula holds for precisely one of the two vertices. In fact, the vertex colours translate directly into logical formulas with one free variable. As a consequence, Colour Refinement distinguishes two graphs if and only if there is a C^2 -sentence that distinguishes them. This correspondence extends to higher dimensions: the k -variable fragment C^k of C corresponds to the $(k - 1)$ -dimensional extension of Colour Refinement, the $(k - 1)$ -dimensional Weisfeiler–Leman algorithm. This algorithm computes a unique colouring for a graph G if and only if G is definable in C^k , i.e. there is a sentence in C^k whose only models are G and its isomorphic copies.

As a matter of fact, the link to the logic C goes even deeper: the number of Colour Refinement iterations required to compute distinct colours corresponds exactly to the quantifier depth of a distinguishing formula. Since the iterations induce a sequence of strictly nested vertex partitions, the process must terminate after at most $n - 1$ rounds, where n is the number of vertices. Consequently, the value $n - 1$ serves as a trivial upper bound on both the number of iterations and the quantifier depth required to distinguish any two (distinguishable) vertices in C^2 .

My talk provides an introduction to the link between the Colour Refinement procedure and the logic C^2 . We revisit a simple characterisation of their expressivity on graphs and on general relational structures. The characterisation implies that the definability of a graph in C^2 can be checked very efficiently.

We then discuss tight lower bounds on the quantifier depth of C^2 -formulas required to distinguish vertices. Through a thorough analysis of computational data from Colour Refinement executions, we constructed infinite families of graphs that witness those bounds. We finish with a presentation of a recent purely theoretical reverse-engineering approach to finding *long-refinement graphs* and a classification of all such graphs with small (or, equivalently, large) degrees.

The talk is based on the collaborations [1, 2, 3] and unpublished work.

2012 ACM Subject Classification Theory of computation → Finite Model Theory; Theory of computation → Graph algorithms analysis; Mathematics of computing → Combinatorial algorithms; Mathematics of computing → Graph theory

Keywords and phrases Colour Refinement, counting logic, Weisfeiler–Leman algorithm, variable width, quantifier depth

Digital Object Identifier 10.4230/LIPIcs.CSL.2026.2

Category Invited Talk



© Sandra Kiefer;

licensed under Creative Commons License CC-BY 4.0

34th EACSL Annual Conference on Computer Science Logic (CSL 2026).

Editors: Stefano Guerrini and Barbara König; Article No. 2; pp. 2:1–2:2

Leibniz International Proceedings in Informatics



LIPICs Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

References

- 1 Sandra Kiefer and T. Devini de Mel. A classification of long-refinement graphs for Colour Refinement. In *Proceedings of the 2026 Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2026*, pages 388–407. doi:10.1137/1.9781611978971.18.
- 2 Sandra Kiefer and Brendan D. McKay. The Iteration Number of Colour Refinement. In *Proceedings of the 47th International Colloquium on Automata, Languages, and Programming, ICALP 2020*, volume 168 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 73:1–73:19, Dagstuhl, Germany, 2020. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.ICALP.2020.73.
- 3 Sandra Kiefer, Pascal Schweitzer, and Erkal Selman. Graphs identified by logics with counting. *ACM Trans. Comput. Log.*, 23(1):1:1–1:31, 2022. doi:10.1145/3417515.