# A category-theoretic approach to automata minimization and learning

Daniela Petrișan
Université Paris Cité, IRIF

In this tutorial we will present generic algorithms for minimization and learning for various forms of automata, starting with deterministic automata, weighted automata and sequential transducers.

We start by describing various forms of automata from a more abstract and unifying perspective. We can think of automata as machines that processes some input, according to its structure (word, tree,...) and produces some effect in some universe of output values (Boolean values, probabilities, weights in a semiring, words over some alphabet,... ). We present a very gentle introduction to some basic notions of category theory required to understand our approach. In particular, we will think of automata as functors from an input category – describing the structure of the inputs – to some output category – describing the computational effect.

We show how minimization results from automata theory can be presented at this level of generality. We identify sufficient conditions for the output category which allow us to obtain an algebraically minimal automaton in a given class. This sheds some light on why minimization works for some classes of automata and why it fails for others.

We then give an overview of Angluin's $L^*$ algorithm and its variations, and see how we can reuse the minimization results in order to obtain a generic learning algorithm for functorial automata.

If time allows we will discuss some recent applications of this framework beyond word automata.

## References

[1] T. Colcombet and D. Petrișan. Automata minimization: a functorial approach. *Log. Methods Comput. Sci.*, 16(1), 2020.

[2] T. Colcombet, D. Petrișan, and R. Stabile. Learning automata and transducers: A categorical approach. In *CSL 2021,*, volume 183 of *LIPIcs*, pages 15:1–15:17. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021.

[3] T. Colcombet and D. Petrișan. Automata in the category of glued vector spaces. In *MFCS 2017*, volume 83 of *LIPIcs*, pages 52:1–52:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2017.

[4] Q. Aristote. Active learning of deterministic transducers with outputs in arbitrary monoids. In *CSL 2024*, volume 288 of *LIPIcs*, pages 11:1–11:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2024.