ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Distributed Computing

Prof. R. Wattenhofer

Data-Centric Generalization in Neural Algorithmic Reasoning

Neural networks can often learn to mimic algorithms on small problem instances but struggle to generalize to larger, harder ones, despite these being trivial for the actual algorithms. While prior work has explored improving architectures or making models more algorithm-aware, this project focuses on the role of data. What kind of training data leads to better generalization? Can we identify or prioritize examples that are more valuable for learning? This project will investigate if we can make use of techniques common in active learning and curriculum learning can help select or structure training data to enable neural models to generalize more reliably to out-of-distribution instances.

The project will focus specifically on problems defined over graphs (e.g., shortest paths, spanning trees, ...) or puzzles, where generalization across different graph sizes and structures is particularly challenging, but correctness is easy to evaluate with the added benefit that new data can be constructed and solved fast and cheap.

The exact scope of the project is not yet fixed and is up for discussion; we are looking for motivated students who are excited to work on the field of Neural Reasoning.

Requirements: Programming skills (Python, C/C++, ...) and knowledge of machine learning. Prior experience on working with graphs or related machine learning frameworks such as Pytorch or Pytorch geometric is an advantage.
We will have weekly meetings to address questions together, discuss progress, and think about future ideas.

Contact

In a few short sentences, please tell us why you are interested in the project and about your coding and machine learning background (i.e., your own projects or courses).

- Joël Mathys: jmathys@ethz.ch, ETZ G63
- Florian Grötschla: fgroetschla@ethz.ch, ETZ G63