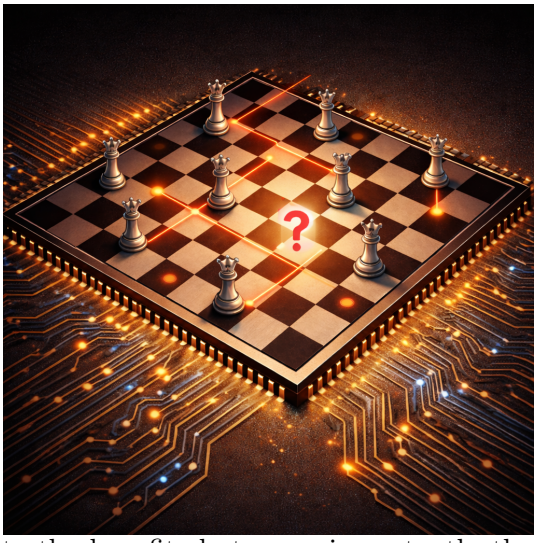
ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Distributed Computing

Prof. R. Wattenhofer

How Small Can Truly Intelligent Networks Be? An Investigation into the Limits of Tiny Recursive Model Reasoning.



Recently proposed recurrent architectures such as HRM or TRM have achieved impressive performance on reasoning benchmarks such as ARC and puzzle solving settings like soduko. Remarkably, they outperform models that are trained on orders of magnitude more data and consist of order of magnitude more parameters. However, many research questions on them still remain unanswered. What is the secret sauce that makes these tiny models work? Can we (theoretically) define in which settings these models are competitive compared to non-recursive, comparable architectures? Do these models suffer from some limitations which have not been highlighted so far?

The aim of the thesis is to properly investigate the benefits but more importantly the hurdles of these models on the way to proper reasoning, ie. can they do well if the path towards a solution cannot always be obtained through purely iterative progress?

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) and knowledge of machine learning. Prior experience on working with related machine learning frameworks such as Pytorch 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
- Giacomo Camposampiero: gcamposampie@ethz.ch, ETZ G63