

# Automated Theorem Proving via Few-Shot Programming

**Advisor:** Jonathan Laurent (jonathan.laurent@kit.edu)

**Keywords:** Theorem Proving, Machine-Learning, Large Language Models, AlphaZero

**Required skills:** good programming skills are expected ; some previous exposure to logic, theorem-proving or program verification would be a plus

## Description:

Large-language models such as GPT4 have proved surprisingly good at solving a wide range of tasks from a handful of examples and at generating structured data such as programs or mathematical formulas. Yet, their lack of reliability puts a limit on their ability to solve large problems that require many steps of reasoning.

Our work proposes a new programming paradigm called “few-shot programming”, where domain experts write high-level problem-solving strategies in the form of nondeterministic programs whose choice-points are annotated with examples. This induces a search space that can be explored by querying large-language models for guidance. In addition, once enough examples are provided by humans for a strategy to bootstrap, this strategy can self-improve via an AlphaZero-like scheme, by solving a series of problems with search and extracting annotations from successful solutions.

We are looking for students to write expert strategies for a variety of tasks in automated theorem-proving, program synthesis and program verification within this paradigm of “few-shot programming”. In particular, doing so successfully will require students to think of original ways to decompose large problems into smaller ones that are tractable for large language models.

**Related published work:** *Learning to Find Proofs and Theorems by Learning to Refine Search Strategies: The Case of Loop Invariant Synthesis*, Jonathan Laurent & André Platzer, NeurIPS 2022