

MARGARIDA FERREIRA

October 2024

✉ margarida@cmu.edu 📁 marghrid.github.io

RESEARCH INTERESTS I am interested in **formal methods** and **program synthesis**, particularly in their application to computer networks. Currently my work focuses on automatically generating correct implementations of stateful programs based on noisy or incomplete specifications.

EDUCATION **PhD. Computer Science** • *Carnegie Mellon University* | *Técnico Lisboa* • 2021 – now
Advised by Professors Ruben Martins and Inês Lynce. Dual Degree PhD fellowship by CMU Portugal with Técnico Lisboa. Expected graduation: 2026.

MSc. Computer Science and Engineering • *Técnico Lisboa* • 2018 – 2020
Advised by Professor Inês Lynce and Dr. Miguel Neves. First year spent at Technical University of Munich, in Germany. Specialization in Artificial Intelligence, Algorithms and Programming. Master thesis: *FOREST: An Interactive Multi-tree Synthesizer for Regular Expressions*, advised by Professor Inês Lynce and Dr. Miguel Neves (OutSystems).

BSc. Computer Science and Engineering • *Técnico Lisboa* • 2015 – 2018
GPA: 18/20. Research project *Satisfying Cooperative Path-finding*, advised by Professor Inês Lynce.

PUBLICATIONS Trace-Guided Synthesis of API-Composing Programs. **M. Ferreira**, V. Nicolet, J. Dodds, D. Kroening. **Under submission.**
Reverse-Engineering Congestion Control Algorithm Behavior. **M. Ferreira**, R. Ware, Y. Kothari, I. Lynce, R. Martins, A. Narayan, J. Sherry. **IMC 2024.**
Counterfeiting Congestion Control Algorithms. **M. Ferreira**, A. Narayan, I. Lynce, R. Martins, J. Sherry. **HotNets 2021.**
FOREST: An Interactive Multi-tree Synthesizer for Regular Expressions. **M. Ferreira**, M. Terra-Neves, M. Ventura, I. Lynce, and R. Martins. **TACAS 2021.**

EXPERIENCE **Applied Scientist Intern** • *Amazon Web Services* • Fall 2024
I am working with the Automated Reasoning for Cloud Operations team, with Joey Dodds and Victor Nicolet, on building an efficient anomaly detection filter for structured logs. We use program synthesis to build a regex- and graph-based ruleset that generalizes the behavior observed in a sequence of non-anomalous logs.

Teaching Assistant

Fall 2023 • Foundations of Programming (undergraduate) • *Técnico Lisboa*.
Fall 2022 • Automated Reasoning and Satisfiability (graduate) • *Carnegie Mellon University*.
Spring 2021 • Algorithms and Data Structures (undergraduate) • *Técnico Lisboa*.
Fall 2020 • Artificial Intelligence (undergraduate) • *Técnico Lisboa*.

Applied Scientist Intern • *Amazon Web Services* • Summer 2023

I worked with the Automated Reasoning for Cloud Operations team, with Joey Dodds and Victor Nicolet, on automatically synthesizing AWS Systems Manager documents, programming scripts that allow users of AWS to define actions to be performed on their managed instances.

Early Stage Researcher • *INESC-ID* • 2020 – 2021

I worked on synthesis of regular expressions from examples, and later on reverse engineering congestion control algorithms, at the Automated Reasoning and Software Reliability group.

Research Intern • *OutSystems* • 2019 – 2020

I worked on synthesis of regular expression form validations, under the supervision of Miguel Neves and Miguel Ventura, at Engineering Department, Artificial Intelligence division.

SKILLS I program mainly in **Python**, including its scientific libraries: **NumPy**, **SciPy**, **Matplotlib**. I have extensive experience with constraint-solving frameworks: **Z3**, **CVC5**, and **Gurobi**, and their APIs. I am also proficient in **C++17** and its standard libraries. I have limited experience with **Rust**, the solver-aided language **Rosette**, and proof assistants **Coq** and **Lean**.

PROJECTS **SYREN** • 2023 – 2024

We propose a new method for synthesizing programs from partial traces of execution. We combine compiler-like optimizing rewrites with programming-by-example to infer program's hidden control- and data-flow. We show the applicability of our method synthesizing cloud automation scripts with different domains.

Jetstream • 2021 – 2022

Jetstream simplifies, restructures, and verifies semantics of virtual switch rulesets. We use SMT to find and remove redundancies from the ruleset, provably maintaining the original semantics. Our analysis of production rulesets shows that JetStream improves packet classification performance by up to 70% using state-of-the-art packet classifiers.

CCA synthesis • 2020 – 2024

We use program synthesis to reverse-engineer congestion control algorithm implementations based on traces collected using the original implementation. Our synthesis procedure uses SMT to encode the space of candidate congestion control algorithms, and tries to find one whose behavior minimizes the error against the given trace. Using this approach, we can synthesize several kernel CCAs including a simplified Reno using only a few traces.

FOREST • 2019 – 2020

FOREST automatically synthesizes regular expressions that match a desired pattern expressed using examples. It uses an SMT solver to explore and prune the search space, and to synthesize capture conditions that ensure the validity of numerical values in the input. We use Z3's regular expressions theory to implement user interaction based on distinguishing inputs. Experimental results show that Forest outperforms previous state-of-the-art regular expression synthesizers.

MISC Article about my work as a PhD summer intern at AWS • Oct 2023

Article about my experience in the CMU Portugal program • May 2022

I helped design and taught the AfterSchool Artificial Intelligence Course by Treetree2, for high school and middle school students who want to learn about Programming and AI • Jan–Feb 2022

FELLOWSHIPS Dual Degree PhD Fellowship • *CMU Portugal* • 2021–26

🏆 **HONORS** Excellent Teacher Award • *Instituto Superior Técnico* • 2020–21

GOLEM Research Scholarship • *INESC-ID, CMU Portugal* • 2019–21

Erasmus+ Scholarship • *European Commission* • 2018–19

New Talents in Artificial Intelligence Scholarship • *Calouste Gulbenkian Foundation* • 2017–18

Academic Excellence Certificates • *Instituto Superior Técnico* • 2015 – 2018, 2020