Pedro Valero

Personal Data

PHONE: +1 650 387 0934
EMAIL: pedro.valero.mejia@gmail.com
WEBPAGE: https://pevalme.github.io/

Work Experience

July 2019 – October 2019 **Facebook** (Palo Alto, California)
Research Intern at the Data Compression Team Manager: *Yann Collet*
We analyzed the strengths of grammar-based compression to understand how it could be used within the next generation of data compressors. In order to perform this analysis, I built a prototype (in *C*) of a grammar-based compressor that achieved compression ratios comparable to the ones obtained with zstd.

September 2016 – Current **IMDEA Software Institute** (Madrid, Spain)
PhD Student Manager: *Pierre Ganty*
My PhD is focused on Applications of Language Theory. Some of the most relevant projects I have worked on as part of my PhD are:

Searching on Compressed Text
We studied the problem of searching with regular expressions on compressed text without decompression. We devised a simple technique that speeds up the search by taking advantage of the information about repetitions on the text extracted by the compressor. We implemented a tool, *zearch*, for counting the number of lines in a compressed file that match a given regular expression. The tool outperforms the state of the art decompress-and-search approach. The results were published at the *Data Compression Conference*.

Automata Minimization
We defined a framework of automata constructions based on equivalences over words that unifies the different existing techniques for automata minimization. The most relevant aspect of this work is that our framework covers the Brzozowski's algorithm, which was previously considered as orthogonal to the rest of minimization algorithms. This work was published at the *Mathematical Foundations of Computer Science* conference.

September 2015 – May 2016 **IMDEA Software Institute** (Madrid, Spain)
Part-time Intern Manager: *Pierre Ganty*
During this internship, we studied whether different idioms, common among network protocols, could be validated with parser generators for context-free languages. We implemented (with *flex* and *bison*) a modular, robust, and efficient input validator for HTTP relying on context-free grammars and regular expressions. We published the obtained results at the *Automated Technology for Verification and Analysis* conference.

June 2015 – September 2015 **Max Planck Institute for Software Systems** (Kaiserslautern, Germany)
Intern Manager: *Rupak Majumdar*
We used the Robot Operative System to simulate a robot which was controlled by a combination of voice commands and hand gestures (captured with a Leap Motion device). I designed and implemented a system to handle the given commands and execute them according to their predefined priorities.

June 2014 – May 2015 **IMDEA Software Institute** (Madrid, Spain)
Intern Manager: *Pierre Ganty*
The goal of this project was to update the tool mist, a safety checker for Petri Nets and extensions developed by my supervisor, Pierre Ganty. In particular, I implemented new *Python* scripts to better test and benchmark the tool and improved the presentation of the results by using the *JavaScript* library *D3*.

Software

HTTValidator	An input validator for HTTP messages that relies on recognizers for context-free and regular languages (implemented using Bison and Flex respectively) to perform the validation. *Publicly available on GitHub.*
Zearch	A tool for regular expression searching on grammar-compressed text (implemented in C). Publicly available on *GitHub*.

Programming Skills

Languages	**Advanced:** C, Python, LaTeX. **Medium**: C++, Java, Bash, Awk, JavaScript, PHP, HTML, CSS. **Basic**: R, SQL, Assembly, Lisp, Prolog.
Software	Linux, Sublime Text, Atom, Git, svn, mercurial, Zsh.

Publications

SAS 2019	Complete Abstractions for Checking Language Inclusion *with Francesco Ranzato and Pierre Ganty.*
MFCS 2019	A Congruence-based Perspective on Automata Minimization Algorithms *with Elena Gutiérrez and Pierre Ganty.*
DCC 2019	Regular Expression Searching on Compressed Text *with Pierre Ganty.*
ATVA 2017	A Language-Theoretic View on Network Protocols *with Pierre Ganty and Boris Köpf.*

Committees

As a PhD student I have contributed to the organization of the ATVA'19 and TACAS'19 conferences as a member of the *Artifact Evaluation Committee*. The goal of these committees is to check consistency and replicability of results presented in submitted papers as well as evaluating their completeness, documentation and ease of use.

Education

2016 - Current	PhD in Software, Systems and Computing at **Universidad Politécnica de Madrid**
2011 - 2016	Double degree at Computer Science and Mathematics at **Universidad Autónoma de Madrid** Obtained four consecutive *Excellence Awards* for academic performance. GPA: 9.14/10.0