

# Pavlos Petoumenos

Informatics Forum, 10 Crichton Street, Edinburgh, EH8 9AB, UK  
ppetoume@inf.ed.ac.uk

## Overview

---

Peer-reviewed journal and conference publications	13
Peer-reviewed workshop publications	8
Citations	264
H-index	8
Papers presented	7
Invited presentations	4
Undergraduate and MSc students co-supervised	8
PhD students co-supervised	4

## Research highlights

---

- **Deep Learning techniques for compiler optimisations** [CGO'17, PACT'17]  
**Automatic benchmark generation** and **heuristic construction** using deep neural networks. Using code from public repositories trains an LSTM network on how real code is constructed. Samples the network to automatically synthesise new code, indistinguishable from real code. Creates a **huge corpus of benchmarks**, orders of magnitude larger than existing corpuses. For training on the benchmarks, it lets the LSTM extract training features from the code. Effectively learning heuristics directly on code. **Better heuristics** with **zero expert effort**. **Best paper award** in CGO'17, already taught in a CS course in Colorado State University. Contribution: Guided the research and originated the ideas of using deep neural networks for code compression and generation.
- **Measuring QoE of interactive workloads** [IISWC'14]  
Fundamentally changes the way we create benchmarks and use them to evaluate performance. Automatic and fast generation of representative benchmarks from real interactive workloads. The first to measure the **user's perception of performance** on arbitrary workloads. Large energy savings by adapting performance to the user's needs. **Best paper award** in IISWC'14. Contribution: Originated the idea, drove and directed the research.
- **Cache replacement based on reuse distance prediction** [ICCD'07, SAMOS'09, JWAC'10]  
Replacement policy emulating the theoretically **optimal policy**. Developed the first hardware mechanism to accurately **predict online data reuse patterns**. **Fifth most cited** cache replacement paper of the last decade with **99 citations**. Contribution: Originated the idea, designed the reuse distance predictors of SAMOS'09 and JWAC'10, and developed the replacement candidate selection algorithm.

## Teaching highlights

---

- **Teaching assistant**

Five years of lab exercises, tutorials, projects, and up to half of the lectures for a fifth year module.

Mostly independent for the first three years, **fully independent** for the last two.

On my own initiative, I **overhauled the syllabus** and replaced all **teaching material**.

- **Programming club and programming contests**

Founder of the University's **programming club** with up to 60 students attending.

Teaches first and second year students basic programming skills.

Each week I choose an algorithmic problem and go through it step by step with the students.

Coach of the University's **programming teams** for the last two years.

Responsible for the selection of the teams, their training, and securing funding.

In NWERC 2016, **18th best team overall, 6th best from the UK**.

## Research grants

---

EPSRC SUMMER project (EP/P003915/1) £120,000.

A heterogeneous OS scheduler for minimal energy consumption which respects the user's perception of performance. Continues the work started with the [IISWC'14].

2016

## Research areas

---

### Active Research

- **Benchmarking methodologies** [IISWC'14, ADAPT'16, CGO'17]: Existing benchmarks in many areas of Computer Science are **too few** and **barely representative** of real applications. My work aims at developing **automatic techniques** for representative **benchmark generation**, targeting mainly mobile and OpenCL applications. This line of research has produced **two best paper awards**.
- **Application tuning on heterogeneous systems** [LCPC'14, ADAPT'16, HLPGPU'16, CGO'17, PACT'17]: Heterogeneous processors make performance tuning far more complicated. More devices to choose from, each one with different optimal runtime and compiler settings, different for each application. My research uses **Machine Learning** to make these complex decisions, **automatically and with little or no expert knowledge**.
- **Fine-grained power profiling** [TACO'17]: Existing on-chip energy measuring subsystems have a granularity in the order of milliseconds. Measurements represent aggregates over millions of instructions, which makes it difficult to pinpoint energy and power inefficiencies. This work developed measurement and analysis methodologies to enable **sub-millisecond resolution** for energy profiling, which in turn makes it possible to apply **fine-grained energy optimisations**.

### Past Research

- **Cache management** [MoBS'06, SAMOS'06, IISWC'06, ICCD'07, SAMOS'09, JWAC'10, CF'10]: Traditional replacement policies, like LRU, do not work well for modern last level caches (LLC) and are even worse for shared LLCs, where they tend to split the cache in unfair ways. My work introduced novel replacement and sharing policies for **near optimal utilisation of the cache** by predicting the locality and reuse patterns of the accessed data. My publications in this area have been cited more than 170 times.
- **MLP-aware Instruction Queue resizing** [ARCS'10]: Much of the energy of out-of-order cores is spent trying to extract Instruction Level Parallelism (ILP). Previous research proposed saving energy by adapting the size of the Instruction Queue (IQ) to the amount of available ILP. My work was the first to show that what determines the optimal resizing is actually the amount of Memory Level Parallelism (MLP). I developed a runtime MLP prediction technique and used it to drive IQ resizing decisions, **reducing its energy consumption by 30%**, with almost no performance impact.
- **Efficient power capping** [IPDPS'09, ICPADS'15]: Processors have to limit performance to keep power consumption within acceptable limits to avoid damaging themselves. Typically this is achieved with DVFS, but with each new technology node DVFS becomes less efficient. My work explores new power capping techniques, as well as more intelligent heuristics to coordinate existing ones.

## Employment history

---

### University of Edinburgh

Edinburgh, UK

RESEARCH ASSOCIATE, EPSRC ALEA AND SUMMER PROJECTS

December 2013 - Present

- Published one journal, five conference, and five workshop papers.
- Authored a successful EPSRC grant proposal.
- Co-supervising **four PhD students** and working with them on a daily basis: Volker Seeker, William Ogilvie, Chris Cummins, and Paschalis Mpeis.
- Chief editor for **Compucast**, a podcast for computer science and co-host for the **Compiler Optimisation Youtube lectures**
- Founder of a **programming club** for undergraduate students, as well as coach for the University's undergraduate **programming teams** (9th and 13th best teams in UKIPC 2016).

## Education

---

### University of Patras

Patras, Greece

PHD IN ELECTRICAL AND COMPUTER ENGINEERING

October 2005 - July 2011

- Supervisor: Associate Professor Stefanos Kaxiras (now Uppsala University).
- Thesis: Managing Shared Resources in Multi-core Systems.
- Published seven conference and three workshop papers.
- Co-supervised the Diploma theses of five students
- Responsible for the tutorials, projects, lab exercises, and part of the lectures of a fifth year module.

### University of Patras

Patras, Greece

DIPLOMA IN ELECTRICAL AND COMPUTER ENGINEERING (5 YEAR DEGREE)

October 2000 - October 2005

- 7.9/10 (top 3% of my year)

## Teaching and mentoring

---

Invited lectures for the "Compiler Optimisations" module, U. of Edinburgh	2016 - Present
Programming club for undergraduate students, U. of Edinburgh	2015 - Present
Co-supervision of four PhD and three MSc students, U. of Edinburgh	2014 - Present
Coach for the University of Edinburgh's programming teams	2014 - Present
Co-supervision of five Diploma projects, U. of Patras	2006 - 2011
Lectures for a fifth year module, Advanced Microprocessors, U. of Patras	2006 - 2011
Teaching assistant for two modules, Microprocessors I and II, U. of Patras	2006 - 2008

## Invited talks and outreach

---

Compiler Optimisations Youtube lectures - Co-host	January 2017 - Present
How fast? How furious? - Lancaster University	October 2016
How fast? How furious? Real optimizations for real people - Glasgow Parallelism Group Seminars, Glasgow	May 2016
Compucast - Chief Editor	March 2016 - Present

<i>How fast? How furious? Real optimizations for real people</i> International Workshop on Architectural and Micro-Architectural Support for Dynamic Optimization (AMAS-DO), Barcelona	March 2016
Iterative Compilation 1.01 - PPar Lunch Lectures, Edinburgh	December 2014

## Professional activities

---

HiPEAC CSW Thematic Session on Big Data Applications, Chair	2016
CADO Workshop, Program Committee	2016
ParCo Conference, Symposium Chair	2015
COSMIC Workshop, General Chair	2014 - Present

## Awards and Scholarships

---

Best Paper Award CGO'17 ( <i>IEEE conference with 25% acceptance rate</i> )	2017
Best Paper Award IISWC'14 ( <i>IEEE conference. Top venue for benchmarking methodologies and characterisation</i> )	2014
HiPEAC collaboration grant ( <i>highly competitive with hundreds of applications every year from PhD students from all around the EU</i> )	2007
Greek State Scholarships Foundation Award for Distinguished Undergraduate Students ( <i>awarded to the top 3 students of each year</i> )	2003

## List of peer-reviewed journal and conference publications

---

Conference and Workshop papers which I have presented have their title in bold.

- [1] *End-to-end Deep Learning of Optimization Heuristics*  
C. Cummins, P. Petoumenos, Z. Wang, and H. Leather. PACT 2017.
- [2] *Synthesizing Benchmarks for Predictive Modeling*  
C. Cummins, P. Petoumenos, Z. Wang, and H. Leather. CGO 2017.  
**Best Paper award**
- [3] ***Minimizing the cost of iterative compilation with active learning***  
W. Ogilvie, P. Petoumenos, Z. Wang, and H. Leather. CGO 2017.
- [4] *ALEA: A Fine-grained Energy Profiling Tool*  
L. Mukhanov, P. Petoumenos, Z. Wang, N. Parasyris, D. Nikolopoulos, B. de Supinski, and H. Leather. TACO 2017.
- [5] ***Power Capping: What Works, What Does Not***  
P. Petoumenos, L. Mukhanov, Z. Wang, H. Leather, and D. Nikolopoulos. ICPADS 2015.
- [6] *Measuring QoE of Interactive Workloads and Characterising Frequency Governors on Mobile Devices*  
V. Seeker, P. Petoumenos, H. Leather, and B. Franke. IISWC 2014.  
**Best paper award**
- [7] *Where replacement algorithms fail: a thorough analysis*  
G. Keramidas, P. Petoumenos, and S. Kaxiras. CF 2010.
- [8] ***MLP-aware Instruction Queue Resizing: The Key to Power-Efficient Performance***  
P. Petoumenos, G. Psychou, S. Kaxiras, J. M. Cebrian Gonzalez, and J. L. Aragon. ARCS 2010.
- [9] ***Instruction-based reuse-distance prediction for effective cache management***  
P. Petoumenos, G. Keramidas, and S. Kaxiras. SAMOS 2009.

- [10] *Efficient microarchitecture policies for accurately adapting to power constraints*  
J. M. Cebrian Gonzalez, J. L. Aragon, J. M. Garcia, P. Petoumenos, and S. Kaxiras. IPDPS 2009.
- [11] **Cache Replacement Based on Reuse-Distance Prediction**  
G. Keramidas, P. Petoumenos, and S. Kaxiras. ICCD 2007.  
**Most cited paper with 99 citations**
- [12] *Using Value Locality to Reduce Memory Encryption Overhead in Embedded Processors*  
G. Keramidas, P. Petoumenos, S. Kaxiras, A. Antonopoulos, and D. Serpanos. ETFA 2007.
- [13] **Modelling Cache Sharing on Chip Multiprocessor Architectures**  
P. Petoumenos, G. Keramidas, H. Zeffer, S. Kaxiras, and E. Hagersten. IISWC 2006.  
**Second most cited paper with 36 citations**

## List of peer-reviewed workshop papers

---

- [W.1] *Towards Collaborative Performance Tuning of Algorithmic Skeletons*  
C. Cummins, P. Petoumenos, M. Steuwer, and H. Leather. HLPGPU 2016.
- [W.2] *Iterative Compilation on Mobile Devices*  
P. Mpeis, P. Petoumenos, and H. Leather. ADAPT 2016.
- [W.3] *Autotuning OpenCL Workgroup Size for Stencil Patterns*  
C. Cummins, P. Petoumenos, M. Steuwer, and H. Leather. ADAPT 2016.
- [W.4] *Intelligent Heuristic Construction with Active Learning*  
W. Ogilvie, P. Petoumenos, Z. Wang, and H. Leather. CPC 2015.
- [W.5] *Fast automatic heuristic construction using active learning*  
W. Ogilvie, P. Petoumenos, Z. Wang, and H. Leather. LCPC 2014.
- [W.6] **Instruction-based Reuse Distance Prediction Replacement Policy**  
P. Petoumenos, G. Keramidas, and S. Kaxiras. JWAC 2010.  
**8th best cache replacement policy out of 25 submissions**
- [W.7] *Preventing Denial-of-Service Attacks in Shared CMP Caches*  
G. Keramidas, P. Petoumenos, S. Kaxiras, A. Antonopoulos, and D. Serpanos. SAMOS 2006.
- [W.8] *Statshare: A statistical model for managing cache sharing via decay*  
P. Petoumenos, G. Keramidas, H. Zeffer, S. Kaxiras, and E. Hagersten. MoBS 2006.