



HAL
open science

A Comprehensive Analysis of Approximate Computing Techniques: From Component- to Application-Level

Alberto Bosio, Daniel Menard, Olivier Sentieys

► To cite this version:

Alberto Bosio, Daniel Menard, Olivier Sentieys. A Comprehensive Analysis of Approximate Computing Techniques: From Component- to Application-Level. ESWEEK 2018 - Embedded Systems Week, Sep 2018, Torino, Italy. pp.1-2. hal-01941755

HAL Id: hal-01941755

<https://inria.hal.science/hal-01941755>

Submitted on 3 Dec 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

TUTORIAL PROPOSAL – ESWEEK'18

A comprehensive analysis of Approximate Computing techniques: from component- to application-level

Abstract:

A new design paradigm, Approximate Computing (AxC), has been established to investigate how computing systems can be more energy efficient, faster, and less complex. Intuitively, instead of performing exact computation and, consequently, requiring a high amount of resources, AxC aims to selectively violate the specifications, trading accuracy off for efficiency. It has been demonstrated in the literature the effectiveness of imprecise computation for both software and hardware components implementing inexact algorithms, showing an inherent resiliency to errors [MIT16], [CHI13].

This tutorial introduces basic and advanced topics on AxC. We intend to follow a bottom-up approach: from component- up to application-level. More in detail, we will first present the main concept and techniques (e.g., functional approximation, voltage over-scaling). We then move to present some compile-time results in terms of energy-efficiency, area, performance versus accuracy of computations when using customized arithmetic (fixed-point, floating-point) and also try to derive some conclusions by comparing the different paradigms [BAR17]. The algorithmic-level approximation methods are then presented. Energy consumption can be reduced by approximating or skipping part of the computation. The concept of incremental refinement, early termination and fast decision will be detailed.

The last part of the tutorial focuses on existing methodology for applying AxC techniques at application-level. Finally, we present as a case study IIDEA [BAR], an automatic framework that compounds (i) a source-to-source manipulation tool and (ii) an evolutionary search engine, in order to find the best functional approximation version of the given C/C++ code.

Intended Audience

Approximate Computing (AxC) is a promising technique for reducing the cost (in terms of area and power consumption) of a computing system. However, AxC is still in its infancy since there is a lack of methodologies and automated tools for the entire design and manufacturing flow. We strongly believe that professionals already working on that will be very interested by the latest advances and experiences presented during the special session. Moreover, the tutorial could be attractive for all the people who want to extend their horizon and learn best practices across the levels and technologies in the areas of AxC.

Syllabus

Tutorial will cover:

1. Introduction
2. AxC techniques: functional approximation, over-scaling approximation
3. Data precision, number representation and approximate arithmetic
4. Algorithmic-level approximation
5. Application-level: Design Space Exploration

We would like to propose a half-day tutorial as lecture style for the first 4 points of the Syllabus. The last part (5) will consist of lab exercises. Attendees are supposed to have their own laptop and basic knowledge of UNIX like environment (bash). No particular SW requirements are demanded since a docker image will be provided.

Speakers:

- Alberto BOSIO – LIRMM (France)
- Daniel MENARD – INSA Rennes (France)
- Olivier SENTIEYS – INRIA (France)

Speaker biographies:

Alberto Bosio received the PhD in Computer Engineering from the Politecnico di Torino, Italy in 2006. From 2007 he is an Associate Professor at LIRMM - University of Montpellier in France. His research interests include Approximate Computing, In-Memory Computing, Test and Diagnosis of Digital circuits and systems and Reliability. He co-authored 1 book, 3 patents 35 journals, and over 120 conference papers. He will be the chair of the ETTTC from January 2018. He is a member of the IEEE.

Web: <http://www.lirmm.fr/~bosio/home/>

Daniel Menard received the Ph.D. and HDR degrees in Signal Processing and Telecommunications from the University of Rennes, respectively in 2002 and 2011. From 2003 to 2012 he was Associate Professor at University of Rennes in France. He is currently Full Professor at INSA Rennes. His research activities focus on the energy efficient implementation of signal and image processing applications in embedded systems. His research topics include approximate computing, fixed-point arithmetic, energy optimization in MPSoC, low power HEVC video encoding and decoding and embedded stereo-vision. He has published 25 papers in international journals and 63 papers in international conferences. He is member of the DISPS Technical Committee of the IEEE Signal Processing Society.

Web: <http://dmenard.perso.insa-rennes.fr>

Olivier Sentieys is a Professor at the University of Rennes and holds an Inria Research Chair on Energy-Efficient Computing Systems. He has more than 20 years of expertise in the fields of system-on-chip architectures, reconfigurable systems and their associated CAD tools, finite arithmetic effects, numerical accuracy analysis and low-power sensor networks. He authored or coauthored more than 150 journal publications or peer-reviewed conference papers and hold 6 patents. In particular, his research on methods for analytical analysis of errors in reduced-precision arithmetic and word-length optimization since 2000 with more than 50 publications, can be considered as a pioneering work in the field of approximate computing.

Web: <http://people.rennes.inria.fr/Olivier.Sentieys/>

References

[MIT16] S. Mittal, "A survey of techniques for approximate computing," *ACM Comput. Surv.*, vol. 48, no. 4, pp. 62:1–62:33, Mar. 2016. [Online]. Available: <http://doi.acm.org/10.1145/2893356>

[CHI13] V. K. Chippa, S. T. Chakradhar, K. Roy, and A. Raghunathan, "Analysis and characterization of inherent application resilience for approximate computing," in *Proceedings of the 50th Annual Design Automation Conference*. ACM, 2013, p. 113.

[BAR] M. Barbareschi, A. Bosio, <http://wpage.unina.it/mario.barbareschi/iideaa/>

[BAR17] B. Barrois, O. Sentieys, and D. Menard, "The Hidden Cost of Functional Approximation Against Careful Data Sizing - A Case Study," In *IEEE/ACM Design Automation and Test in Europe (DATE)*, page 6, 2017.