Guest Editors' Introduction

Hellebrand, Sybille; Henkel, Jörg; Raghunathan, Anand;
Wunderlich, Hans-Joachim

IEEE Embedded Systems Letters Vol. 10(1), March 2018

doi: https://doi.org/10.1109/LES.2018.2789942

Abstract: Approximate computing exploits the inherent error resilience of many applications to optimize power consumption, run time, or chip area. Especially in audio, image, and video processing, but also in data mining or resource allocation tasks, approximate results are "good enough" for many application domains and may be hard to distinguish from perfect results. But even in domains where accurate results are required, approximate computing can be applied in various layers of the computing stack. Approximate computing promises to significantly increase computing efficiency (especially performance per power) and power density and is therefore of particular interest for embedded computing where these metrics are key.

Preprint

Guest Editors' Introduction

Approximate Computing exploits the inherent error resilience of many applications to optimize power consumption, run time, or chip area. Especially in audio, image and video processing, but also in data mining or resource allocation tasks, approximate results are "good enough" for many application domains and may be hard to distinguish from perfect results. But even in domains where accurate results are required, approximate computing can be applied in various layers of the computing stack. Approximate Computing promises to significantly increase computing efficiency (especially performance per power) and power density and is therefore of particular interest for embedded computing where these metrics are key.

Approximate Computing is a multi-faceted area of research with many approaches spanning hardware and software. This is also reflected by the contributions in this special issue. The first paper by Thierry Moreau, Joshua San Miguel, Mark Wyse, James Bornholt, Armin Alaghi, Luis Ceze, Natalie Enright Jerger and Adrian Sampson introduces the reader to this highly active research area by classifying existing techniques with respect to their correctability, reproducibility and controllability. A deeper discussion of different characteristics in these three dimensions further provides valuable input for further research in approximate computing. In the next paper, Minho Ha and Sunggu Lee address approximation at the hardware level and describe a component for processing partial results in multiplication. Compared to previous solutions, the presented 4:2 compressor ensures that deviations from the exact result are always of the same type, which simplifies corrections whenever needed. Another contribution by Samira Ataei and James E. Stine is dedicated to a new voltage-scalable SRAM architecture working with three different supply voltages to improve static noise margin during read and write modes and to reduce leakage current in retention mode. The SRAM cells are realized in such a way, that the number of unprotected approximate bits in a memory word can easily be adapted to the requirements of a specific application. As shown by Mohsen Imani, Daniel Peroni and Tajana Rosing in the following paper, appropriate memory support is also beneficial for the acceleration of computing tasks. Here, a special CAM, which supports a fast and energy-efficient search for entries with best match, is used as a lookup table for the results of frequent computations. The computing accuracy can be tuned by controlling the use of the approximate lookup table with the help of a weighted similarity metric.

Approximation-aware high-level synthesis is in the focus of the next contribution. Seogoo Lee and Andreas Gerstlauer suggest clustering loop iterations according to data dependent requirements. This way a good compromise can be achieved between a coarse-grained processing of the overall loop and a fine-grained processing of individual loop iterations. In contrast, the next paper by Tobias Isenberg, Marie-Christine Jakobs, Felix Pauck and Heike Wehrheim investigates the robustness of software verification in the context of approximate computing. The presented approach extracts constraints during the verification process that can guide approximate hardware selection or design. Finally, in the last contribution of this special issue Ilia Polian discusses test and reliability challenges in approximate computing. Despite the inherent error tolerance, approximate computing may not be robust against arbitrary hardware defects. Starting from existing techniques like threshold testing, which have been originally developed for yield improvement, the concepts of error magnitude testing and error rate testing are introduced before general reliability issues are discussed. On the one hand, approximate computing may also tolerate some hardware defects, but on the other hand hidden hardware defects may evolve into hard failures over time and cause severe problems.

In summary, we believe that the contributions of this special issue reflect the current state-of-the-art, while also identifying challenges that are yet to be solved in the field of approximate computing.

Sybille Hellebrand,
Paderborn University, Germany

Jörg Henkel,
Karlsruhe Institute of Technology, Germany

Anand Raghunathan,
Purdue University, USA

Hans-Joachim Wunderlich
University of Stuttgart, Germany