UCLA Computer Science Department: CS 259 Prof. Miloš D. Ercegovac

Current Topics in Computer Science:
Arithmetic in the Design of Neural Networks

Engineering 6 468B
825-5414; e-mail: milos@cs.ucla.edu
http://www.cs.ucla.edu/˜milos

Spring 2019, MW 10:00 - 11:55 am, Boelter 4413
Office hours: M2:00 - 3:00pm or by appointment

Course material: Lecture viewgraphs, papers, handouts, etc. posted on **CCLE Course.**
Recommended Textbook: M.D. Ercegovac and T. Lang, *Digital Arithmetic*, Morgan and Kaufmann, 2004.
Related resources: Books, journals, conferences, Web sites, simulators .
Grading: Based on three design explorations, reports, project, presentations, and participation in class discussions.

SEMINAR DESCRIPTION

Introduction. There is a renewal of interest in digital arithmetic due to demanding computational needs of neural networks in machine learning. Over the years research in digital arithmetic successfully enabled high performance in general-purpose processors and in domain-specific systems such as digital signal processors. The research has been a synergism between number systems and arithmetic algorithms compatible with VLSI design, with a goal of fast, reliable, and energy-efficient standardized operations compatible with established instruction-set architectures. With the rise of domain-specific accelerators, and massive computing systems such as neural networks, standardizing arithmetic operations is no longer desirable. This opens opportunities and challenges in developing new arithmetic solutions. The objective of this research seminar is understanding of ideas and techniques to find such solutions. Our study will focus on unconventional approaches in representation, arithmetic algorithms – in particular, left-to-right instead of right-to-left processing of digits, and approximate arithmetic.

Example of a research question we will be exploring: 1. We know how to perform all operations in left-to-right manner, where the result digit z_j, $j = 1, ..., n$ is obtained after receiving the input digit $x_{j+\delta}$, where δ is a small integer. This allows successive operations to be performed every δ cycles, independent of precision n. 2. What is the speedup in deep recurrent networks?. 3. What is the reduction in complexity of interconnections and memory bandwidth? 4. What is the effect on energy consumption?

Activity A. We study several unconventional arithmetic number representations (such as redundant systems, residue number systems, logarithmic systems, and the recent POSIT floating-point system). These representations have properties potentially useful in the design of massive arithmetic networks. Second, we identify several types of algorithms that are relevant in the design of neural networks. We consider both conventional and unconventional approaches and compare their features. By far the most frequent operation is multiply-add. It is used in matrix operations and in computing inner and tensor products. Low-precision and variable precision formats are gaining in popularity, with a mix of fixed-point and floating-point arithmetic. It is possible to do multiplications in a left-to-right (MS digit first) manner which eliminates the need for a final adder. Moreover, MSDF mode allows efficient overlap of multiplications and accumulation of products. A generalization of the MSDF mode is online arithmetic where in all operations the inputs are consumed and results are obtained digit-by-digit starting with the most-significant digits. Such a mode allows digit-level overlap between successive operations, reducing delays due to data dependencies, and enabling massive parallelism. The required bandwidth of interconnections is reduced to single digit per operand/result simplifying the design and reducing the energy. Moreover, on-line arithmetic supports gracefully variable-precision operations without the need for rounding and multiple designs. Importantly, online mode allows zero-delay max/min operations. These operational features of on-line arithmetic will be investigated in realizing arithmetic operations and composite algorithms for inner and tensor products, convolution, sum of squares and Euclidean norms, gradient descent computations, softmax function, exponentials and logarithms, hyperbolic tangent, linear and logistic regressions, and low-precision linear algebra. These topics will be covered and discussed in lectures.

Activity B. We will study selected papers on different types of neural networks (NNs). These include feedforward NNs, convolutional (CNN), recurrent (RNN), recursive NNs, LSTM NNs, and binarized neural networks (BNN).

Activity C. You will work on three **design explorations** focusing on arithmetic aspects in neural networks. In each design exploration your main goal will be to determine tradeoffs in the design and try to find optimal arithmetic solutions. You will summarize your work in a report and make presentations.

Project. One of the successful design explorations you may expand into your **final project**. Alternatively, you may choose a topic not related to design explorations.

The activities will be overlapped.