TO: The Faculty of the College of Engineering

FROM: Elmore Family School of Electrical and Computer Engineering

RE: New Graduate Course, ECE 60827 Programmable Accelerator Architectures

The faculty of the School of Electrical and Computer Engineering has approved the following new course. This action is now submitted to the Engineering Faculty with a recommendation for approval.

ECE 60827 Programmable Accelerator Architectures
Sem. 2, Lecture 3, Cr. 3.
Prerequisite by Topic: ECE 43700 or equivalent, ECE 56500

Description: Programmable hardware accelerators seek to fulfill the promise of continued performance and energy-efficiency gains in the era of a slowing Moore's law, larger problem sizes and an increased focused on energy-efficiency. These factors have caused hardware acceleration to become ubiquitous in today's computing world and critically important in computing's future. This class will introduce students to the architectures of programmable accelerators. We will delve deeply into the architectures of modern massively parallel accelerators like GPUs, culminating in a course project. General topics in hardware acceleration will be discussed, including but not limited to GPGPU and massively parallel computing, approximate accelerators, reconfigurable hardware and programmable hardware for machine learning.

Reason: Programmable hardware accelerators seek to fulfill the promise of continued performance and energy-efficiency gains in the era of a slowing Moore's law, larger problem sizes and an increased focused on energy-efficiency. These factors have caused hardware acceleration to become ubiquitous in today's computing world and critically important in computing's future.

Milind Kulkarni,
Associate Head for Teaching and Learning
Elmore Family School of Electrical and Computer Engineering

ECE 695 – Programmable Accelerator Architectures

EDX Spring 2021

Overview

Programmable hardware accelerators seek to fulfill the promise of continued performance and energy-efficiency gains in the era of a slowing Moore's law, larger problem sizes and an increased focused on energy-efficiency. These factors have caused hardware acceleration to become ubiquitous in today's computing world and critically important in computing's future.

This class will introduce students to the architectures of programmable accelerators. We will delve deeply into the architectures of modern massively parallel accelerators like GPUs, culminating in a course project. General topics in hardware acceleration will be discussed, including but not limited to GPGPU and massively parallel computing, approximate accelerators, reconfigurable hardware and programmable hardware for machine learning.

Textbooks

Aamodt, T.M., Fung, W.L., & Rogers, T.G. 2018. General Purpose Graphics Processor Architecture. Synthesis Lectures on Computer Architecture.
Kirk, D. B., & Hwu, W. M. W. (2016). Programming Massively Parallel Processors: A Hands-on Approach: Third Edition. Elsevier Inc. (Recommended)

Grading Scheme

Quizzes:	25%
Midterm Exam:	20%
Programming Assignments:	30%
Final Project:	25%

Letter Grades:

- A+: 97-100
- A: 93-96
- A-: 90-92
- B+: 87-89
- B: 83-86
- B-: 80-82
- C+: 77-79
- C: 73-76
- C-: 70-72
- D+: 67-69
- D: 63-66
- D-: 60-62
- F: 59 or below

Paper Reading

We will read papers at a rate of approximately 1/week. See Piazza for list of currently assigned papers. Papers critiques will be submitted via blackboard. See paper critique guidelines for a list of expectations for paper reviews.

Assignments

We will have 1-2 programming assignments through the semester. Assignment #1 will be a CUDA programming assignment intended to familiarize you with the software aspects of the class. Assignment #2 may be a simulator-based programming assignment. These Assignments should be done individually.

Exams

There will be one midterm.

Project

There is a final group project due in the last week of class. Students will be expected to give presentations on their project to the class during the final week(s) of class. The exact nature of the project will be up to the student. We will begin the project selection process after ~1 month of class.

Rough Course Outline

1 Week	General purpose architecture background and the evolution to accelerators
2 Weeks	Programming massively parallel accelerators
2 Weeks	Advances in the GPU programming model. Introduction to pervasive ML algorithms (i.e. CNNs)
2 Weeks	GPU core design
3 Weeks	GPU memory system and interconnect
2 Weeks	CPU/GPU systems and AMD Fusion architecture
3 Weeks	Case studies on architectures for machine learning

Addendum: EFD 9-22

Previous Course Offerings:

ECE 69500 Programmable Accelerator Architecture, Spring 21 (14), Spring 20 (36), Spring 18 (21)