Electrical and Electronics Engineering Institute
University of the Philippines, Diliman, Quezon City

CoE 163 Computing Architecture and Algorithms

Study Guide: Week 5

(April 19-23, 2021)

Introduction

This week we will focus on optimizing matrix-matrix multiplication (MMM), which is a fundamental linear algebra operation that is essential to many scientific computations in many applications. Speeding up and improving the efficiency of MMM will help improve the performance of these scientific algorithms.

There are many ways to optimize MMM. We focus on only a few in this course, but the discussion should help you appreciate other optimization methods on your own if you are interested to study these on your own. The material first discusses how memory behaves depending on how the loops of an MMM algorithm are set up. Then we discuss how to use blocking in MMM to improve the behavior of the cache in our algorithm.

The detailed schedule for the week is as follows:

Time (min)	Activity	Topic	Assessment	Score
10	📦	Study guide		
15	📦	Review of Cache Organization and Behavior		
20	📦 🖥	Optimizing MMM Part 1: Impact of traversal/loop order		
30			Assignment	10
20	📦 🖥	Optimizing MMM Part 2: MMM with Blocking		
20			Online quiz	10

Objectives

The objectives for this week's activities are for you to be able to:

- explain how traversal by row or column in an MMM algorithm loop can impact performance
- Explain how blocking improves performance of MMM

Activity 1: Review of Cache Behavior

Go through the slides [Week5a: Cache Behevior Review](). This is simply a review of cache organization and behavior that you've covered in previous courses. It is important for you to visualize the cache clearly so that you can understand the succeeding slides.

Activity 2: Optimizing MMM Part 1: Impact of traversal/loop order

After recalling how a computer cache works, go through the slides entitled [Week 5b: Matrix-Matrix Multiplication (Part 1)](). You can reinforce your understanding of this section by checking out the Cache Simulator[1] linked in UVLe to see an example of good and bad cache behavior when traversing a 2D array. There will be an optional Zoom session where I will discuss the slides on **May 6, 2021, 10:30am**. Details are given below and also linked on UVLe. The session will be uploaded to Youtube so that you can view it at your own time afterwards.

Meeting ID: 980 0311 4064
Password: 568643

Join Zoom Meeting
[https://prime-ph.zoom.us/j/98003114064?pwd=VnE4ZlZTbzdzUWhma1diRDFpc1cwdz09]()

An assignment will be given to let you try benchmarking various 2D array traversals and compare their performance. Please await instructions on UVLe.

Activity 3: Optimizing MMM Part 2: Matrix-Matrix Multiplication with Blocking

The last part of this week's topic introduces another way to optimize MMM (and potentially other matrix algorithms) by using a method called blocking. Go

---

[1] http://vhosts.eecs.umich.edu/370simulators/cache/simulator.html

through the slides Week 5c: Matrix-Matrix Multiplication (Part 2) that discusses this. For additional understanding of how blocking works, watch the video MMM with blocking - How does it work?[2]. If you are able to attend the synchronous session linked above, Part 2 will also be covered. The Youtube link will be posted on UVLe after the Zoom session so that you can view it at your own time.

A knowledge check quiz will be opened on UVLe to assess what you've learned from this week's lesson. Please check UVLe and await announcements for this quiz.

End of Week 5
Next Week's Topic: Linear Algebra Software Libraries and Continuation of Matrix Operations

---

[2] https://www.youtube.com/watch?v=HJDc7MkpvEc