ECE 3803, GPU Programming for General Computing
Syllabus

Jeff Epstein

Fall 2024

Contents

1 Basic information

1.1 Course description

This course provides a foundation for development, analysis, and optimization of general programs that take advantage of the computational capabilities of modern GPU hardware.

1.2 Learning outcomes

Upon successful completion of this course, students should be able to:

- Create programs that effectively utilize modern GPU architectures
- Analyze the performance characteristics of GPU programs
- Optimize the performance of GPU programs based on metrics and architectural considerations
- Implement algorithms as GPU programs

ABET student outcomes:

- an ability to identify, formulate, and solve complex engineering problems by applying principles of engineering, science, and mathematics

1.3 Logistics

Meeting times are as follow:

Section	Time	Place	Instructor
Lecture	3:30pm-4:45pm MW	Van Leer C341	Epstein

Please attend only the section for which you are registered.

1.4 Contact

The principal instructor for this course is:

Name	email	Office hours	Office hours location
Jeff Epstein	jeff.epstein@gatech.edu	by appointment	Klaus 3308

The TAs for this course are:

Name	email	Office hours	Office hours location
Owen P Kew	okew3@gatech.edu	5pm-6pm M, 1pm-2pm W	Klaus 1446/1448
Aamay M Puntambekar	apuntambekar6@gatech.edu	2pm-3pm TTh	Zoom link

Office hours begin the second week of class.

In addition to the scheduled office hours specified above, the course instructor offers office hours by appointment. Please request an appointment by email.

When contacting a member of the teaching staff by email, please use the course code ("ECE 3803") in the subject, so that we know the context.

Please do not send source code as an email attachment. Assignments are not accepted by email.

1.5 Prerequisites

Students should have completed ECE 2035, ECE 2036, or equivalent.

In particular, students should be familiar with:

- programming in the C language or the C++ language,
- the design of microprocessors, including datapath, memory, and instruction set.

1.6 Textbook

Our textbook is Programming Massively Parallel Processors: A Hands-on Approach, 4th Edition, 2022, by Hwu, Kirk, and El Hajj. This textbook is optional.

1.7 Topics and schedule

Anticipated topics to be covered include the following.

- Introduction and context
  - review of CPU architecture
  - limits of sequential programming
  - review of C programming
- GPU execution
  - CUDA compilation
  - data parallelism
  - parallel programming
  - race conditions
- Memory organization and access
  - GPU memory hierarchy
  - memory management
  - shared memory
  - bank conflicts
  - memory coalescing
- Performance optimization
  - occupancy
  - profiling
- Parallel patterns
  - atomics
  - reduction
  - scanning
  - sorting

- – synchronization
- GPU microarchitecture
  - – example
  - – assembly

Coursework may include programming assignments focused on GPU applications in audio/video processing, linear algebra, and machine learning.

We present a tentative course schedule. Topics and dates are subject to change without notice.

Week	Date	Note	Topic	Assignment
1	19 Aug 2024		syllabus, overview	
	21 Aug 2024		what is a GPU	
2	26 Aug 2024		review of C	hw0
	28 Aug 2024		CUDA intro, error checking	
3	2 Sep 2024	LABOR DAY	NO CLASS	
	4 Sep 2024		GPU memory hierarchy	
4	9 Sep 2024		shared memory, race conditions	
	11 Sep 2024		bank conflicts	
5	16 Sep 2024		boundary conditions	
	18 Sep 2024		memory coalescing	
6	23 Sep 2024		L1 cache	
	25 Sep 2024		occupancy	
7	30 Sep 2024		divergence	
	2 Oct 2024		divergence	
8	7 Oct 2024		midterm review	
	9 Oct 2024		midterm exam	
9	14 Oct 2024	FALL BREAK	NO CLASS	
	16 Oct 2024		atomics	
10	21 Oct 2024		reduction	
	23 Oct 2024		scan	
11	28 Oct 2024		profiling	
	30 Oct 2024		sorting	
12	4 Nov 2024		libraries	
	6 Nov 2024		streams	
13	11 Nov 2024		synchronization	
	13 Nov 2024		multi-GPU	
14	18 Nov 2024		GPU microarchitecture	
	20 Nov 2024		PTX assembly	
15	25 Nov 2024			
	27 Nov 2024	TURKEY BREAK	NO CLASS	
16	2 Dec 2024		final review	
	4 Dec 2024	READING DAY	NO CLASS	
	6 Dec 2024	2:40pm-5:30pm	final exam	

2 Coursework

2.1 Grading rubric

Your final numeric grade will be calculated with the following composition:

Component	Weight
Homework	40%
Midterm exam	30%
Final exam	30%

Your final numeric grade, expressed as a percent of the maximum possible numeric grade, will be converted to a final letter grade based on the following table:

Letter	A	B	C	D	F
Minimum	90	80	70	60	0

2.2 Homework

Homework is intended to let you practice your new skills and knowledge. Homework will be posted on Canvas and will be submitted on the Gradescope system.

Each homework has the same weight: your overall homework grade is the average of all your homework grades.

2.3 Exams

Exams represent an opportunity to prove your new skills and knowledge. Exams are administered during class time. Students who are not present in class during the exam will receive a grade of zero for that exam. During the exam, use of outside resources, use of phones, communication between students, and use of any other program or web site is prohibited.

Exams are to be taken on your own personal computer, which you must bring with you to class on days when an exam is scheduled. If you anticipate not being able to bring your computer to class, please contact your instructor at least one week in advance, and we will find an alternative solution. If you do not have your computer on the day of a exam and have not already made alternative plans with your instructor, you will receive a grade of zero for that exam.

In order to get credit for an exam, you must be physically present in class during the administration of the exam, and must present your Georgia Tech ID card to a proctor upon request. Failure to comply will result in receiving a grade of zero for that exam. Attempting to circumvent these policies may constitute a violation of academic integrity.

3 Technology and tools

Each student should have consistent access to a personal desktop or laptop computer with either Linux, Mac OS, or Windows operating system, and a stable internet connection.

Students are responsible for maintaining their computing equipment in good working condition at all times. Technical issues with personal computer equipment and internet connection are not valid reasons for requesting deadline extensions or other accommodations.

Students are responsible for protecting against data loss. It is recommended that students create regular backups of their work. Losing work due to accident or malfunction is not a valid reason for requesting deadline extensions or other accommodations.

Students are responsible for ensuring access to the Georgia Tech network infrastructure and services, using the Georgia Tech VPN if necessary.

3.1   Canvas

We will use Canvas for course information and discussion.

Students are encouraged to make use of Canvas's discussion forum to ask and answer questions about course material. The principal instructor, the TAs, and other students will do their best to help you.

A one-point final grade bonus is offered to the five students with the highest number of helpful answers and/or insightful questions on the Canvas discussion forum.

Although you are free to use the discussion system to discuss course material, posting solutions is prohibited by the course's academic integrity policy. Posting inappropriate comments on the discussion forum may result in a grade penalty.

3.2   Gradescope

Homework will be submitted on Gradescope. Gradescope is accessible through Canvas.

You can submit your work as many times as you want before the deadline. All work must be submitted before the deadline. You must verify that all necessary files have been submitted before the deadline. After the deadline, you will not the have opportunity to add or change submitted files.

After your work has been graded, you'll be able to review your grade on Gradescope. Students should carefully review their grade to make sure they understand any applied penalties and to learn from the feedback. In addition, you should review your graded work in a timely manner if you wish to submit a regrade request.

3.3   GT Instructional Cluster

The GT Instructional Cluster (ICE) provides us access to GPU hardware and related software. We will use this resource for compiling and running programs. Please refer to the document *A Guide to the GT Instructional Cluster*, available on Canvas, for information about using the cluster.

3.4   CUDA Toolkit

We will use Nvidia's CUDA Toolkit to develop software for GPUs. It's available on the GT Instructional Cluster.

If your computer has an Nvidia GPU and a compatible operating system, you can install the toolkit on your computer. This will allow you to compile CUDA programs locally. You can download the CUDA Toolkit from the official download link. If you are using Linux, you may be able to find it in your distribution's software repository.

Please note that the course's teaching staff will not provide support with installing or using the CUDA Toolkit on your own computer. The only officially supported use of the Toolkit is on the GT Instructional Cluster.

3.5 Nvidia Documentation

There is extensive official documentation provided by Nvidia about their GPUs and how to develop software for them. Students are expected to refer to this documentation as necessary. Here are some helpful links:

- CUDA C Programming Guide
- CUDA Compiler
- PTX Instruction Set Architecture

4 Policies

4.1 Classroom expectations

Attendance is in-person. Remote attendance via Zoom is not offered. Lectures will not be recorded. Students are expected to arrive to every class promptly: please avoid causing a disruption with a late arrival. You should be actively engaged in the learning process during the duration of the class time.

In the event of a campus closure due to inclement weather, lectures will be delivered remotely, in accordance with Georgia Tech policies.

You are expected to take notes during lecture. Please be prepared to do so.

Active participation in class discussions is strongly encouraged. This is the best time for students to ask questions or clarify any confusing concepts. To participate, please just raise your hand. Students are responsible for any material covered in class, even if it isn't in the textbook. If you miss a class, you should contact a classmate to recover the missed content and assignments.

When in class, please disable or silence any device that may audibly disrupt the class.

During class time, please do not use your phone or computer for activities unrelated to the course. Please do not eat, drink, or sleep during class.

4.2 Communication

Your instructor may use email to make announcements. It is your responsibility to check your university email account regularly.

Assignments will typically be posted on Canvas. It is your responsibility to check Canvas for assignments, and to submit your work in a timely manner.

4.3   Deadlines, absences, make-ups, extra credit, late submissions

A principle of fairness in evaluation is that all students receive the same opportunity to complete the same work. This principle informs our class policy that the schedule of assignments will not be modified for individual students. In particular:

- All deadlines are strict, final, and automatically enforced. It is the student's responsibility to ensure that they submit their assignments well in advance of the posted deadline. If you miss the deadline for any reason, the submission will not be accepted. Please do not wait until the last minute!

- In general, there is no opportunity for students to make up a missed assessment (such as an exam or quiz), or to reschedule such an assessment. Please ensure that you are present during assessments.

- Assignments are given to the class as a whole, not to individual students. It would be unfair to offer an extra-credit assignment to an individual if others in the class do not have the same opportunity. Please do not ask for such an unfair advantage.

- Although it is recognized that occasionally it may be necessary for students to be absent from scheduled classes or laboratories for personal reasons, students are responsible for all material covered in their absences, and they are responsible for the academic consequences of their absences.

However, make-up or rescheduling opportunities may be allowed in certain exceptional situations:

- In the event of a family emergency; or a medical emergency or an illness that is severe enough to require medical attention, students are responsible for contacting the Office of the Dean of Students as soon as possible to report the medical issue or emergency, providing dated documentation from a medical professional and requesting assistance in notifying their instructors. The medical documentation will be handled confidentially within the Office of the Dean of Students and will inform a decision as to whether communication with instructional faculty is appropriate.

- Students who are absent because of participation in approved Institute activities (such as field trips, professional conferences, and athletic events) will be permitted to make up the work missed during their absences. Approval of such activities will be granted by the Student Academic and Financial Affairs Committee of the Academic Senate, and statements of the approved absence may be obtained from the Office of the Registrar. To request approval, please fill out the Institute-Approved Absence form at least 2 weeks prior to the event.

- Students who are absent because of participation in a particular religious observance will be permitted to make up the work missed during their absence with no late penalty, provided the student informs the course instructor of the upcoming absence, in writing, within the first two weeks of class, and provided the student makes up the missed material within the time frame established by the course instructor.

More information is provided in the relevant section of the course catalog.

4.4 Academic accommodations

Georgia Tech offers accommodations to students with disabilities. If you anticipate issues related to the format or requirements of this course, please meet with the instructor within the first week of the semester to discuss ways to ensure your full participation in the course. If you determine that disability related accommodations are necessary, please register with the Office of Disability Services as soon as possible, and present your accommodation letter to the instructor. Students who have already been approved for accommodations through the Office of Disability Services, please schedule an appointment with your instructor to discuss your academic accommodations. Afterwards, accommodations must be requested from the instructor one week in advance of each instance of their use.

4.5 Regrades

You are entitled to a fair and accurate assessment of your work. If you believe that your work has been graded unfairly or inaccurately, you may request a secondary evaluation (henceforth, "regrade") of your work from your instructor.

A request for regrade must be submitted to your instructor by email within one week of the initial release of the grade. A regrade request must specify the assignment and question to be considered, and must contain an explanation of why the student believes that the grading is unfair or inaccurate. Regrade requests are not accepted by means other than email. Regrade requests received after one week will not be accepted, and at that time all grades are considered final: students are advised to check their grades and review their work promptly. Grades released at the end of the semester may have a shorter window for regrades.

Please note that a regrade may result in a revised grade higher, lower, or the same as the original grade. Submitting regrade requests in bad faith or without appropriate justification will result in a grade penalty.

4.6 Getting help

This is a challenging course. If you are struggling with the material, it is your responsibility to take action to improve. There are a variety of ways to get help:

- Ask questions in class. A classroom discussion is the most direct way to understand the material thoroughly.
- Meet with a TA. The teaching assistants offer regular office hours.
- Meet with the instructor. The instructor has regular office hours, as well as office hours by appointment.
- Send an email. You can contact your instructor or a TA by email to ask questions about the homework or course material.
- Read the book. You may prefer the presentation of material in a textual format.
- Ask on the discussion forum. The class site may have an on-line discussion forum where you can post questions.

When asking for help, please consider these guidelines:

- **Don't wait until the last minute.** It's your responsibility to complete assignments on time. Plan accordingly: allow enough time to get any requisite help *before* the deadline.

- **Provide information along with your question.** The person helping you will need to know something about your problem. Be sure to tell them what you are trying to do, what you already did, what happened, and what you expected to happen. Include any relevant code, error messages, or screen shots. Don't just say "It didn't work."
- **Don't expect someone else to do your work for you.** The goal of asking for help is to gain a better understanding of course material. However, when it comes to graded assignments, you still need to do the work for yourself. Expect guidance, not solutions.

4.7 Academic integrity

We aim to ensure that your grade reflects your understanding of the material, and therefore we require that all work that you submit, including but not limited to any homework, quiz, exam, or project, be a result solely of your own understanding and effort, except for ungraded in-class exercises, and for those submitted works where collaboration among students is explicitly allowed.

This course is bound by Georgia Tech's Student Code of Conduct and Honor Code. Suspected violations of academic integrity may be handled through Faculty Conference Resolution or via referral to the Office of Student Integrity.

Motivation

Why do we take academic integrity seriously?

- By committing academic dishonesty, you are depriving yourself of an opportunity to learn. The challenging work in this course is intended to enhance your skill and understanding. If you bypass that work, you will have failed to learn the material and reduced the value of your college education. Your grade has value only if it accurately represents what you've accomplished to achieve it.

- Academic dishonesty damages the university as an institution. Employers preferentially hire Georgia Tech graduates because the Georgia Tech degree is conferred on those who have achieved a certain level of education. Acquiring a degree through dishonest means damages the university's reputation and brand.

- Finally, academic dishonesty produces an environment that is unfair to your classmates. If you achieve a grade through dishonest work, students who complete the assignments honestly are at a relative disadvantage.

Policy

We realize that it's not always clear what activities are acceptable in the context of working on your assignments. This is especially true in our field, where the web can provide excellent learning resources (good!) as well as solutions to assignments (bad!).

To help you understand the boundaries of what is and is not acceptable, we provide the following guidelines, which are based on the Stanford CS107 collaboration rules. We classify activities into three categories, according to a traffic light model: those activities that are always acceptable (green light), those that are acceptable but require a citation (orange light), and those that are never acceptable (red light).

Acceptable activities

- Green light Discussion of general course topics.

  You may freely discuss the course material, outside of the context of a specific assignment. This means you can ask and answer questions about the rules of the programming languages and tools that we use, as well as theoretical matters that were covered in lecture or in the textbook. However, the discussion may not refer to solutions for a specific assignment.

  For example: "What does the `elem` function do? What is the syntax for a function definition?"

- Green light Discussion of assignment requirements.

  You may freely discuss the requirements for a specific assignment. However, the discussion may not refer to solutions.

  For example: "Are we allowed to use `reverse`? Is performance important?"

- Green light Use of public resources for background information.

  Web sites, books, and other public resources provide lots of great information. You are expected and encouraged to refer to documentation for the languages and tools that we use, as long as they don't refer to the specific task for your assignment.

  For example, the official Python documentation is okay, but "How to write Space Invaders in Python" is not okay (assuming the assignment is to write Space Invaders).

- Green light Discussion with teaching staff.

  This is the best way to get help! You can always turn to your instructor and TAs with any questions. You can ask questions about any assignment during class, during office hours, or by email.

Activities requiring a citation   What is a citation? If you participate in one of the activities in this section, you must state so in the assignment that it pertains to. If the assignment is a coding exercise, you must write a comment in your code describing the activity. If the assignment is non-code work, you must write a preface describing the activity. Appropriate citation of all external sources is required, including the acknowledgment of any collaboration or assistance.

Every citation must describe (a) the specific source of the information that you used (a person, a book, or a web site), and (b) how you used that information, and to what extent it influenced your work. Failing to provide a citation when participating in one of these activities renders the activity unacceptable, and therefore a violation of academic integrity.

- Orange light Re-using your own work.

  If you are submitting your own work, including code, that you originally completed before you took this course, you must provide a note stating this fact, including when you originally did the work and under what conditions.

  If you want to submit work that you have previously completed for an earlier instance of this course, or for any other course, you must discuss the matter with your instructor before proceeding. In general, you are not allowed to re-submit work completed in an earlier semester.

An exception to this rule applies if you have previously violated academic integrity in this course. In that case, you are prohibited from re-submitting any violating work or any work based on that work.

- Orange light Discussion of testing.

  If you aren't sure if your solution is correct, it can be helpful to develop a strategy to test it. You may propose and discuss tests collaboratively, as long as you provide a note describing the incident. For example, you may remind a friend that a particular function needs to work with empty lists. However, you may not see each other's code or discuss it in detail.

- Orange light Discussion of debugging.

  You may discuss particular error messages and propose possible solutions, as long as you provide a note describing the incident. For example, if a friend says they are getting a "segmentation fault," you may suggest that they might be dereferencing a null pointer. However, you may not see each other's code or discuss it in detail.

Unacceptable activities

- Red light Copying code or answers from others.

  The work you submit must be written entirely by you. This means that the ideas behind it are yours. You may not copy code, read code, or consult code from another person, directly or indirectly. If you are in an environment where you are able to see another person's solution, it is your responsibility to take yourself out of that environment and actively prevent the possibility of being influenced by their work.

- Red light Collaborative development.

  You may not work together with another person, and your submitted solution must not be influenced by anyone else's input. Any sharing of technique, design, pseudocode, structure, or substance, even in partial or incomplete form, is not acceptable. Collaboration of any kind is prohibited: you must work on your assignments alone, from beginning to end. Use of tutors or tutoring services on your assignments and assessments is prohibited.

- Red light Using public resources for assignment-specific code.

  Web sites, books, and other public resources can provide valuable information, but any resource that refers to the task that you are to complete for the assignment, or to a task that is substantially similar to it, is not allowed. The design and implementation of your solution may not be influenced or inspired by an external resource. Taking code, in any quantity, from a web site is prohibited. If you find yourself reading a web site that addresses the design or implementation of a program similar to the task at hand, you are required to stop using it.

  In particular, consulting "coursework aid sites" such as Chegg and CourseHero is prohibited.

  The boundary between "general background resources" and "assignment-specific resources" may be fuzzy. If you have any doubt that a resource you are using is acceptable, ask your instructor first.

- Red light Allowing others to use your code or answers.

At no point should anyone else have access to your work. This means that you may not keep your work on a shared or public computer. You may not post your work on a public web site (including Github and similar sites). You may not show your work to others, even if they promise not to use it in their own work. Your work should remain private, and if at any time it isn't, then you have contributed to academic dishonesty.

You may use Github (and similar sites) as version control storage, but it is your responsibility to ensure that your repository is private. If your code is publicly visible at any time, it constitutes a violation of this policy.

Posting your work on "coursework aid sites" such as Chegg and CourseHero is prohibited.

The prohibition on sharing your work remains in effect even after you have submitted your work, after the due date has passed, and after the course has ended. Be aware that other students may attempt to submit your work as their own without your knowledge. Students in future semesters should not have access to your work.

- Red light Developing code or solutions with the help of artificial intelligence.

  Your submitted work must be based solely on your own understanding. Completing your work with the help of an AI system, such as ChatGPT or Copilot, trained on other people's work violates that condition. Therefore, using these or any similar system is prohibited.

- Red light Violating the environmental conditions for an assessment (such as a quiz or exam).

  Assessments are typically administered in a proctored environment, where students do not have access to external resources. Violating the requirements for the administration of an assessment by, for example, communicating with other people during the time period of the assessment, tampering with or bypassing any software tools used in the administration of the assessment, using prohibited materials or devices during the time period of the assessment, or accessing the assessment remotely, represent a violation of academic integrity.

  Examination materials are available only to those students taking the assessment during the period of the assessment. Exfiltrating the content of an assessment (for example, by photographing questions) is a violation of academic integrity.

- Red light Deliberately providing false information.

  Falsifying or failing to provide a required citation is unacceptable. Questions about the provenance of work must be answered honestly. Attempting to deceive the instructor is a violation of course policy.

Additional academic integrity policies

If you are in doubt about whether a particular activity is acceptable, it's best to consult with your instructor *before* doing it.

If you are struggling in the course, please discuss your situation with your instructor. Violating academic integrity is not a solution.

Students may be interviewed about any material they submit for grading. Students are expected to understand and to be able to explain any code they submit. If you are unable to explain your code, then you probably didn't write it.

Violations of academic integrity will result in immediate referral to the Office of Student Integrity. Penalties will depend on the severity of the transgression and each individual student's history of transgressions. Penalties range from a zero score on an assignment or test, failing the course, or even more severe university-wide actions such as suspension or expulsion from the university.

Preventing violations of academic integrity is everyone's responsibility. If you become aware of an instance of academic dishonesty, even if you are not directly involved in it, you are obligated to report it to your instructor.

If you have violated academic integrity, your first step should be to reach out to your instructor so we can work together to find a way for you to move forward with integrity. A timely request to retract a compromised work or apply an amended citation that properly credits authorship could rectify the misrepresentation that would have otherwise been the basis for a strict punishment.

4.8   Professionalism and respect

At Georgia Tech, we believe that it is important to strive for an atmosphere of mutual respect, acknowledgment, and responsibility among faculty members and the student body. See the course catalog for an articulation of some expectations.

All members of the university community should treat each other with respect. Inappropriate behavior in or out of the classroom is not acceptable. If you experience or witness inappropriate treatment, please contact the instructor.

Violations of this policy may warrant a grade penalty or a referral to the Office of Student Integrity.

4.9   Student well-being

The Center for Mental Health Care & Resources offers in-person and virtual appointments. For more information on these and other student services, contact Student Life.