



Силабус курсу

ПРОЕКТУВАННЯ КОМП'ЮТЕРНИХ СИСТЕМ НА ОСНОВІ ГРАФІЧНИХ ПРОЦЕСОРІВ

Освітньо-професійна програма «Комп'ютерна інженерія»

Ступінь вищої освіти – бакалавр

Спеціальність: 123 «Комп'ютерна інженерія»

Рік навчання: 4, Семестр: 8

Кількість кредитів: 5, Мова викладання: українська

Керівник курсу

ПП

к.т.н. Петро ЛЯЩИНСЬКИЙ

Контактна інформація p.liashchynskyi@gmail.com

Опис дисципліни

Метою курсу «Проектування комп'ютерних систем на основі графічних процесорів» є отримання знань та навиків розробки програмних систем, які призначені для обробки з допомогою графічних процесорів. В процесі роботи студенти отримають знання щодо принципів обробки даних з допомогою центрального та графічного процесорів для розуміння принципу роботи сучасних систем. Додатково студенти отримують знання щодо підходів до розробки розпаралелених алгоритмів та отримають практичні навички для програмної реалізації.

Структура курсу

№ п/п	Тема	Результати навчання	Завдання
1	Вступ. Поняття складності обчислювальних задач	Класифікація обчислювальних задач. Приклади складних обчислювальних задач. Визначення обчислювальної складності задач. Часові та просторові обмеження	Питання
2	Підходи до декомпозиції складних задач	Метод декомпозиції задач. Розбиття задач на підзадачі. Принципи модульного підходу. Визначення граничних умов	Питання, лабораторна робота
3	Підходи до виконання складних програм	Паралельні та послідовні програми. Стратегії ефективного виконання програм. Використання багатоядерних систем. Особливості багатопотокових систем	Питання, лабораторна робота
4	Технологія CUDA: Вступ і архітектура	Вступ до CUDA. Архітектура CUDA. Роль графічних процесорів у паралельних обчисленнях. Переваги використання GPU	Питання, лабораторна робота
5	Вступ до моделей паралельних обчислень	Принципи паралельних обчислень. Загальна архітектура паралельних систем. Основні поняття багато процесорних систем. Переваги та обмеження паралелізму	Питання, лабораторна робота

6	Види розпаралелення обчислювальних задач	Поняття паралелізму. Рівні розпаралелення: блоки, потоки. Типи паралельних задач. Масштабування обчислювальних процесів	Питання, лабораторна робота
7	Синхронізація, потоки та процеси	Визначення потоків та процесів. Основні підходи до синхронізації. Засоби для синхронізації процесів. Приклади використання потоків	Питання, лабораторна робота
8	Складність обчислень паралельних задач	Оцінка складності паралельних алгоритмів. Часові та ресурсні обмеження. Співвідношення між кількістю процесів і швидкістю виконання. Блокування та конкуренція потоків	Питання, лабораторна робота
9	Багатопоточність: основи та бібліотеки	Моделі багатопоточності. Огляд бібліотек для багатопоточних систем. Взаємодія між потоками в багатопотокових програмах. Приклади багатопоточних програм	Питання, лабораторна робота
10	Паралелізм на рівні даних	Знати основи паралельних обчислень з використанням даних, паралельні структури даних. Приклади паралелізму на рівні даних. Оптимізація розподілу даних	Питання, лабораторна робота
11	CUDA: Ядро та функції	Ядро CUDA. Оголошення функцій у CUDA. Взаємодія хоста і девайса. Стратегії для ефективної роботи з функціями	Питання, лабораторна робота
12	CUDA: Потоки та типи пам'яті	Організація потоків у CUDA. Різні типи пам'яті в CUDA. Приклади роботи з потоками. Оптимізація використання пам'яті в GPU	Питання, лабораторна робота
13	CUDA: Блоки та потоки	Структура блоків і потоків у CUDA. Програмування ідентифікаторів блоків і потоків. Взаємодія між блоками. Паралелізм на рівні блоків	Питання, лабораторна робота
14	CUDA: Атомарні операції та синхронізація	Основи атомарних операцій у CUDA. Важливість синхронізації потоків. Підходи до синхронізації на рівні GPU. Приклади застосування атомарних операцій.	Питання, лабораторна робота
15	CUDA: Оптимізація обчислень	Методи оптимізації програм на базі CUDA. Визначення вузьких місць в обчислювальних процесах. Стратегії оптимізації пам'яті. Приклади практичної оптимізації.	Питання, лабораторна робота
16	CUDA: Обробка зображень	Методи обробки зображень на GPU. Впровадження фільтрів і алгоритмів обробки. Прискорення обробки зображень за допомогою CUDA. Приклади з реальних задач	Питання, лабораторна робота
17	Алгоритми сортування даних з допомогою CUDA	Типи алгоритмів сортування. Використання CUDA для швидкого сортування. Оптимізація процесу сортування на GPU. Приклади практичних задач.	Питання, лабораторна робота
18	Робота з нейронними мережами	Нейронні мережі та GPU. Використання CUDA для навчання нейронних мереж. Прискорення обчислень у нейронних мережах за допомогою GPU. Огляд бібліотек для нейронних мереж	Питання, лабораторна робота

19	CUDA: Використання для обробки великих даних	Особливості паралельної обробки великих обсягів даних на GPU. Оптимізація завдань великої вибірки з використанням CUDA. Методи зберігання та доступу до великих даних у CUDA. Приклади обробки великих даних у реальних проектах	Питання, лабораторна робота
20	Практичне застосування CUDA у задачах штучного інтелекту	Впровадження обчислень для AI задач на базі CUDA. Оптимізація нейронних мереж для GPU. Розробка програмного забезпечення для AI з використанням GPU. Використання CUDA у реальних AI задачах	Питання, лабораторна робота

Літературні джерела

1. Березький О.М. Методи, алгоритми та програмні засоби опрацювання біомедичних зображень: монографія / [О.М. Березький, Ю.М. Батько, К.М. Березька, С.О. Вербовий, Т.В. Дацко, Л.О. Дубчак, І.В. Ігнатєв, Г.М. Мельник, В.Д. Николіук, О.Й. Піцун]; під наук. ред. Березький О.М., Тернопіль . ТНЕУ «Економічна думка», 2017. - 330 с
2. Berezsky O. GPU – based biomedical image processing / O. Berezsky, O. Pitsun, L. Dubchak, P. Lyaschynsky, P. Lyaschynsky // Proceedings of XIV International Conference Perspective Technologies and methods in mems design (MEMSTECH 2018) 18-22 April, 2018, Lviv-Polyana, Ukraine, pp. 96- 99
3. Ryan N Gutenkunst, dadi.CUDA: Accelerating Population Genetics Inference with Graphics Processing Units, Molecular Biology and Evolution, Volume 38, Issue 5, May 2021, Pages 2177–2178.
4. Duane Storti. CUDA for Engineers: An Introduction to High-Performance Parallel Computing 1st Edition/ Addison-Wesley Professional; 1st edition (November 2, 2015) – 452 p.
5. Xu, Derong, Y. Luo, C. Montag, Y. Hao, and J. Qiang. Full Range Tune Scan Studies Using Graphics Processing Units With CUDA in EIC beam-beam simulations. No. BNL-222292-2021-COPA. Brookhaven National Lab.(BNL), Upton, NY (United States), 2021..
6. Brian Tuomanen Hands-On GPU Programming with Python and CUDA: Explore high-performance parallel computing with CUDA 1st Edition, Kindle Edition – 2018
7. Jaegeun Han, Bharatkumar Sharma, Learn CUDA Programming - Packt Publishing (September 27, 2019) – 508p.
8. Bhaumik Vaidya. Hands-On GPU-Accelerated Computer Vision with OpenCV and CUDA: Effective techniques for processing complex image data in real time using GPUs Kindle Edition - Packt Publishing; 1st edition (September 26, 2018) – 382 p.
9. Pitsun O. Multi-threaded Parallelization of Automatic Immunohistochemical Image Segmentation. In: Hu, Z., Wang, Y., He, M. (eds) Advances in Intelligent Systems, Computer Science and Digital Economics IV. CSDEIS 2022. Lecture Notes on Data Engineering and Communications Technologies, vol 158. Springer, Cham
10. Aslam, Maria, Omer Riaz, Shahzad Mumtaz, and Ali Daniyal Asif. "Performance comparison of GPU-based jacobi solvers using CUDA provided synchronization methods." IEEE Access 8 (2020): 31792-31812.

Політика оцінювання

Політика щодо дедлайнів і перескладання. Для виконання індивідуальних завдань і проведення контрольних заходів встановлюються конкретні терміни. Перескладання

модулів відбувається з дозволу дирекції факультету за наявності поважних причин (наприклад, лікарняний).

Політика щодо академічної доброчесності. Використання друкованих і електронних джерел інформації під час контрольних заходів та екзаменів заборонено.

Політика щодо відвідування. Пропуски практичних занять обов'язково відпрацьовуються в години консультацій, в іншому випадку вони вважаються оцінкою «0» та враховуються при визначенні середнього арифметичного. Для здобувачів, які навчаються за індивідуальним графіком, поточне оцінювання проводиться під час консультацій, та шляхом виконання завдань в системі Moodle. За об'єктивних причин (наприклад, карантин, воєнний стан, хвороба, закордонне стажування) навчання може відбуватись в онлайн формі за погодженням із керівником курсу з дозволу дирекції факультету.

Оцінювання

Підсумковий бал (за 100-бальною шкалою) з дисципліни «Комп'ютерна логіка» визначається як середньозважена величина, в залежності від питомої ваги кожної складової залікового кредиту:

Модуль 1		Модуль 2		Модуль 3	Модуль 4
10 %	10 %	10 %	20 %	10 %	40%
Поточне оцінювання	Модульний контроль 1	Поточне оцінювання	Модульний контроль 1	Тренінги	Екзамен
Середнє арифметичне за 2 лабораторних роботи	Тестові завдання	Середнє арифметичне за 2 лабораторних роботи	Письмова робота: 2 теоретичних питання, 1 задача, тестові завдання	Виконання завдання	2 теоретичних питання 2 по 25 бали = 50 балів, Задача = 50 балів

Шкала оцінювання:

За шкалою ЗУНУ	За національною шкалою	За шкалою ECTS
90-100	Відмінно	A (відмінно)
85-89	Добре	B (дуже добре)
75-84		C (добре)
65-74	Задовільно	D (задовільно)
60-64		E (достатньо)
35-59	Незадовільно	FX (незадовільно, з можливістю повторного складання)
1-34		F (незадовільно, з обов'язковим повторним курсом)