

## Professional Cuda C Programming

The CUDA Handbook begins where CUDA by Example (Addison-Wesley, 2011) leaves off, discussing CUDA hardware and software in greater detail and covering both CUDA 5.0 and Kepler. Every CUDA developer, from the casual to the most sophisticated, will find something here of interest and immediate usefulness. Newer CUDA developers will see the hardware processes commands and how the driver checks progress; more experienced CUDA developers will appreciate the expert coverage of topics such as the driver API and context migration, as well as the guidance on how best to structure CPU/GPU data interchange and synchronization. The accompanying open source code—more than 25,000 lines of it, freely available at [www.cudahandbook.com](http://www.cudahandbook.com)—is specifically intended to be reused and repurposed by developers. Designed to be both a comprehensive reference and a practical cookbook, the text is divided into the following three parts: Part I, Overview, gives high-level descriptions of the hardware and software that make CUDA possible. Part II, Details, provides thorough descriptions of every aspect of CUDA, including Memory Streams and even Models of execution, including the dynamic parallelism feature, new with CUDA 5.0 and SM 3.5 The streaming multiprocessors, including descriptions of all features through SM 3.5 Programming multiple GPUs Texturing The source code accompanying Part II is presented as reusable microbenchmarks and microdemos, designed to expose specific hardware characteristics or highlight specific use cases. Part III, Select Applications, details specific families of CUDA applications and key parallel algorithms, including Streaming workloads Reduction Parallel prefix sum (Scan) N-body Image Processing These algorithms cover the full range of potential CUDA applications.

This updated and expanded second edition of the Professional CUDA C Programming provides a user-friendly introduction to the subject Taking a clear structural framework, it guides the reader through the subject's core elements. A flowing writing style combines with the use of illustrations and diagrams throughout the text to ensure the reader understands even the most complex of concepts. This succinct and enlightening overview is a required reading for all those interested in the subject. We hope you find this book useful in shaping your future career & Business.

Unlike many C programming books written by C programmers, this brief, self-teaching introduction was written by an instructor familiar with the needs of students. The book defines key programming terms as it teaches the basics of programming. It contains numerous real world programming examples showing first the algorithm, immediately followed by the program for the algorithm, and then its output. End of chapter exercises with "hints" help to review and master material under discussion. An appendix with fifteen "C Lab projects" with their solutions is also included. Features: \*

- \* Defines key programming terms as it teaches the C programming language
- \* Covers major topics such as arrays and pointers, structures and unions, file handling, and more
- \* Includes numerous real world programming examples showing first the algorithm, followed by the program itself, then the desired output

Explore GPU-enabled programmable environment for machine learning, scientific applications, and gaming using PuCUDA, PyOpenGL, and Anaconda Accelerate Key Features Understand effective synchronization strategies for faster processing using GPUs Write parallel processing scripts with PyCuda and PyOpenCL Learn to use the CUDA libraries like CuDNN for deep learning on GPUs Book Description GPUs are proving to be excellent general purpose-parallel computing solutions for high performance tasks such as deep learning and scientific computing. This book will be your guide to getting started with GPU computing. It will start with introducing GPU computing and explain the architecture programming models for GPUs. You will learn, by example, how to perform GPU programming with Python, and you'll look at using integrations such as PyCUDA, PyOpenCL, CuPy and Numba with Anaconda for various tasks such as machine learning and data mining. Going further, you will get to grips with GPU work flows, management, and deployment using modern containerization solutions. Toward the end of the book, you will get familiar with the principles of distributed computing for training machine learning models and enhancing efficiency and performance. By the end of this book, you will be able to set up a GPU ecosystem for running complex applications and data models that demand great processing capabilities, and be able to efficiently manage memory to compute your application effectively and quickly. What you will learn Utilize Python libraries and frameworks for GPU acceleration Set up a GPU-enabled programmable machine learning environment on your system with Anaconda Deploy your machine learning system on cloud containers with illustrated examples Explore PyCUDA and PyOpenCL and compare them with platforms such as CUDA, OpenCL and ROCm. Perform data mining tasks with machine learning models on GPUs Extend your knowledge of GPU computing in scientific applications Who this book is for Data Scientist, Machine Learning enthusiasts and professionals who wants to get started with GPU computation and perform the complex tasks with low-latency. Intermediate knowledge of Python programming is assumed.

Game Over or Next Level?

A Comprehensive Guide to GPU Programming

Hands-On GPU-Accelerated Computer Vision with OpenCV and CUDA

Hands-On GPU Programming with CUDA

An Integrated Approach

Explore the capabilities of GPUs for solving high performance computational problems

*Parallel and High Performance Computing offers techniques guaranteed to boost your code's effectiveness. Summary Complex calculations, like training deep learning models or running large-scale simulations, can take an extremely long time. Efficient parallel programming can save hours—or even days—of computing time. Parallel and High Performance Computing shows you how to deliver faster run-times, greater scalability, and increased energy efficiency to your programs by mastering parallel techniques for multicore processor and GPU hardware. About the technology Write fast, powerful, energy efficient programs that scale to tackle huge volumes of data. Using parallel programming, your code spreads data processing tasks across multiple CPUs for radically better performance. With a little help, you can create software that maximizes both speed and efficiency. About the book Parallel and High Performance Computing offers techniques guaranteed to boost your code's effectiveness. You'll learn to evaluate hardware*

architectures and work with industry standard tools such as OpenMP and MPI. You'll master the data structures and algorithms best suited for high performance computing and learn techniques that save energy on handheld devices. You'll even run a massive tsunami simulation across a bank of GPUs. What's inside Planning a new parallel project Understanding differences in CPU and GPU architecture Addressing underperforming kernels and loops Managing applications with batch scheduling About the reader For experienced programmers proficient with a high-performance computing language like C, C++, or Fortran. About the author Robert Robey works at Los Alamos National Laboratory and has been active in the field of parallel computing for over 30 years. Yuliana Zamora is currently a PhD student and Siebel Scholar at the University of Chicago, and has lectured on programming modern hardware at numerous national conferences. Table of Contents PART 1 INTRODUCTION TO PARALLEL COMPUTING 1 Why parallel computing? 2 Planning for parallelization 3 Performance limits and profiling 4 Data design and performance models 5 Parallel algorithms and patterns PART 2 CPU: THE PARALLEL WORKHORSE 6 Vectorization: FLOPs for free 7 OpenMP that performs 8 MPI: The parallel backbone PART 3 GPUS: BUILT TO ACCELERATE 9 GPU architectures and concepts 10 GPU programming model 11 Directive-based GPU programming 12 GPU languages: Getting down to basics 13 GPU profiling and tools PART 4 HIGH PERFORMANCE COMPUTING ECOSYSTEMS 14 Affinity: Truce with the kernel 15 Batch schedulers: Bringing order to chaos 16 File operations for a parallel world 17 Tools and resources for better code Discover a variety of data-mining algorithms that are useful for selecting small sets of important features from among unwieldy masses of candidates, or extracting useful features from measured variables. As a serious data miner you will often be faced with thousands of candidate features for your prediction or classification application, with most of the features being of little or no value. You'll know that many of these features may be useful only in combination with certain other features while being practically worthless alone or in combination with most others. Some features may have enormous predictive power, but only within a small, specialized area of the feature space. The problems that plague modern data miners are endless. This book helps you solve this problem by presenting modern feature selection techniques and the code to implement them. Some of these techniques are: Forward selection component analysis Local feature selection Linking features and a target with a hidden Markov model Improvements on traditional stepwise selection Nominal-to-ordinal conversion All algorithms are intuitively justified and supported by the relevant equations and explanatory material. The author also presents and explains complete, highly commented source code. The example code is in C++ and CUDA C but Python or other code can be substituted; the algorithm is important, not the code that's used to write it. What You Will Learn Combine principal component analysis with forward and backward stepwise selection to identify a compact subset of a large collection of variables that captures the maximum possible variation within the entire set. Identify features that may have predictive power over only a small subset of the feature domain. Such features can be profitably used by modern predictive models but may be missed by other feature selection methods. Find an underlying hidden Markov model that controls the distributions of feature variables and the target simultaneously. The memory inherent in this method is especially valuable in high-noise applications such as prediction of financial markets. Improve traditional stepwise selection in three ways: examine a collection of 'best-so-far' feature sets; test candidate features for inclusion with cross validation to automatically and effectively limit model complexity; and at each step estimate the probability that our results so far could be just the product of random good luck. We also estimate the probability that the improvement obtained by adding a new variable could have been just good luck. Take a potentially valuable nominal variable (a category or class membership) that is unsuitable for input to a prediction model, and assign to each category a sensible numeric value that can be used as a model input. Who This Book Is For Intermediate to advanced data science programmers and analysts. C++ and CUDA C experience is highly recommended. However, this book can be used as a framework using other languages such as Python.

Summary Manning's bestselling Java 8 book has been revised for Java 9! In *Modern Java in Action*, you'll build on your existing Java language skills with the newest features and techniques. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the Technology Modern applications take advantage of innovative designs, including microservices, reactive architectures, and streaming data. Modern Java features like lambdas, streams, and the long-awaited Java Module System make implementing these designs significantly easier. It's time to upgrade your skills and meet these challenges head on! About the Book *Modern Java in Action* connects new features of the Java language with their practical applications. Using crystal-clear examples and careful attention to detail, this book respects your time. It will help you expand your existing knowledge of core Java as you master modern additions like the Streams API and the Java Module System, explore new approaches to concurrency, and learn how functional concepts can help you write code that's easier to read and maintain. What's inside Thoroughly revised edition of Manning's bestselling *Java 8 in Action* New features in Java 8, Java 9, and beyond Streaming data and reactive programming The Java Module System About the Reader Written for developers familiar with core Java features. About the Author Raoul-Gabriel Urma is CEO of Cambridge Spark. Mario Fusco is a senior software engineer at Red Hat. Alan Mycroft is a University of Cambridge computer science professor; he cofounded the Raspberry Pi Foundation. Table of Contents PART 1 - FUNDAMENTALS Java 8, 9, 10, and 11: what's happening? Passing code with behavior parameterization Lambda expressions PART 2 - FUNCTIONAL-STYLE DATA PROCESSING WITH STREAMS Introducing streams Working with streams Collecting data with streams Parallel data processing and performance PART 3 - EFFECTIVE PROGRAMMING WITH STREAMS AND LAMBIDAS Collection API enhancements Refactoring, testing, and debugging Domain-specific languages using lambdas PART 4 - EVERYDAY JAVA Using Optional as a better alternative to null New Date and Time API Default methods The Java Module System PART 5 - ENHANCED JAVA CONCURRENCY Concepts behind CompletableFuture and reactive programming CompletableFuture: composable asynchronous programming Reactive programming PART 6 - FUNCTIONAL PROGRAMMING AND FUTURE JAVA EVOLUTION Thinking functionally Functional programming techniques Blending OOP and FP: Comparing Java and Scala Conclusions and where next for Java CUDA Fortran for Scientists and Engineers shows how high-performance application developers can leverage the power of GPUs using Fortran, the familiar language of scientific computing and supercomputer performance benchmarking. The authors presume no prior parallel computing experience, and cover the basics along with best practices for efficient GPU computing using CUDA Fortran. To help you add CUDA Fortran to existing Fortran codes, the book explains how to understand the target GPU architecture, identify computationally intensive parts of the code, and modify the code to manage the data and parallelism and optimize performance. All of this is done in Fortran, without having to rewrite in another language. Each concept is illustrated with actual examples so you can immediately evaluate the performance of your code in comparison. Leverage the power of GPU computing with PGI's CUDA Fortran compiler Gain insights from members of the CUDA Fortran language development team Includes multi-GPU programming in CUDA Fortran, covering both peer-to-peer and message passing interface (MPI) approaches Includes full source code for all the examples and several case studies Download source code and slides from the book's companion website

GPU Parallel Program Development Using CUDA

Heterogeneous Computing with OpenCL 2.0

Parallel and High Performance Computing

From Algorithms to Programming on State-of-the-Art Platforms

Data Parallel C++

Algorithmic Puzzles

An overview of the most prominent contemporary parallel processing programming models, written in a unique tutorial style. With the coming of the parallel computing era, computer scientists have turned their attention to designing programming models that are suited for high-performance parallel computing and supercomputing systems. Programming parallel systems is complicated by the fact that multiple processing units are simultaneously computing and moving data. This book offers an overview of some of the most prominent parallel programming models used in high-performance computing and supercomputing systems today. The chapters describe the programming models in a unique tutorial style rather than using the formal approach taken in the research literature. The aim is to

cover a wide range of parallel programming models, enabling the reader to understand what each has to offer. The book begins with a description of the Message Passing Interface (MPI), the most common parallel programming model for distributed memory computing. It goes on to cover one-sided communication models, ranging from low-level runtime libraries (GASNet, OpenSHMEM) to high-level programming models (UPC, GA, Chapel); task-oriented programming models (Charm++, ADLB, Scioto, Swift, CnC) that allow users to describe their computation and data units as tasks so that the runtime system can manage computation and data movement as necessary; and parallel programming models intended for on-node parallelism in the context of multicore architecture or attached accelerators (OpenMP, Cilk Plus, TBB, CUDA, OpenCL). The book will be a valuable resource for graduate students, researchers, and any scientist who works with data sets and large computations. Contributors Timothy Armstrong, Michael G. Burke, Ralph Butler, Bradford L. Chamberlain, Sunita Chandrasekaran, Barbara Chapman, Jeff Daily, James Dinan, Deepak Eachempati, Ian T. Foster, William D. Gropp, Paul Hargrove, Wen-mei Hwu, Nikhil Jain, Laxmikant Kale, David Kirk, Kath Knobe, Ariram Krishnamoorthy, Jeffery A. Kuehn, Alexey Kukanov, Charles E. Leiserson, Jonathan Lifflander, Ewing Lusk, Tim Mattson, Bruce Palmer, Steven C. Pieper, Stephen W. Poole, Arch D. Robison, Frank Schlimbach, Rajeev Thakur, Abhinav Vishnu, Justin M. Wozniak, Michael Wilde, Kathy Yelick, Yili Zheng

Capitalize on the faster GPU processors in today's computers with the C++ AMP code library—and bring massive parallelism to your project. With this practical book, experienced C++ developers will learn parallel programming fundamentals with C++ AMP through detailed examples, code snippets, and case studies. Learn the advantages of parallelism and get best practices for harnessing this technology in your applications. Discover how to: Gain greater code performance using graphics processing units (GPUs) Choose accelerators that enable you to write code for GPUs Apply thread tiles, tile barriers, and tile static memory Debug C++ AMP code with Microsoft Visual Studio Use profiling tools to track the performance of your code

**Programming Massively Parallel Processors: A Hands-on Approach, Second Edition**, teaches students how to program massively parallel processors. It offers a detailed discussion of various techniques for constructing parallel programs. Case studies are used to demonstrate the development process, which begins with computational thinking and ends with effective and efficient parallel programs. This guide shows both student and professional alike the basic concepts of parallel programming and GPU architecture. Topics of performance, floating-point format, parallel patterns, and dynamic parallelism are covered in depth. This revised edition contains more parallel programming examples, commonly-used libraries such as Thrust, and explanations of the latest tools. It also provides new coverage of CUDA 5.0, improved performance, enhanced development tools, increased hardware support, and more; increased coverage of related technology, OpenCL and new material on algorithm patterns, GPU clusters, host programming, and data parallelism; and two new case studies (on MRI reconstruction and molecular visualization) that explore the latest applications of CUDA and GPUs for scientific research and high-performance computing. This book should be a valuable resource for advanced students, software engineers, programmers, and hardware engineers. New coverage of CUDA 5.0, improved performance, enhanced development tools, increased hardware support, and more Increased coverage of related technology, OpenCL and new material on algorithm patterns, GPU clusters, host programming, and data parallelism Two new case studies (on MRI reconstruction and molecular visualization) explore the latest applications of CUDA and GPUs for scientific research and high-performance computing

Machine generated contents note: 1. How to think in CUDA 2. Tools to build, debug and profile 3. The GPU performance envelope 4. The CUDA memory subsystems 5. Exploiting the CUDA execution grid 6. MultiGPU applications and scaling 7. Numerical CUDA, libraries and high-level language bindings 8. Mixing CUDA with rendering 9. High Performance Machine Learning 10. Scientific Visualization 11. Multimedia with OpenCV 12. Ultra Low-power Devices: Tegra.

**Hands-On GPU Computing with Python**

**An Introduction to General-Purpose GPU Programming, Portable Documents**

**Best Practices for Efficient CUDA Fortran Programming**

**CUDA Fortran for Scientists and Engineers**

**CUDA by Example**

**The CUDA Handbook**

Developers who want to access USB devices from their embedded systems will find a helpful resource in **USB Embedded Hosts: The Developer's Guide**. This new book from the author of **USB Complete** shows how small systems can take advantage of the same wealth of USB devices available to conventional PCs. The book begins with a review of USB host communication protocols. Readers then learn which USB host requirements are relaxed for embedded systems and what new requirements some embedded systems must meet. To help in selecting a development platform, the book explores available hardware and software for USB host communications in small systems. The heart of the book focuses on communicating with USB devices. The topics (with example code) include USB drives, keyboards, virtual serial ports, network bridges, mics, speakers, video cameras, and printers, plus devices that don't fit defined USB classes. Also discussed are systems that support both USB host and device functions. The example code is written for the BeagleBoard-xM open development board using a distribution of Linux targeted to small systems. Also covered is how to use Linux commands and utilities to learn about, monitor, and debug communications with USB devices.

Many of today's complex scientific applications now require a vast amount of computational power. General purpose graphics processing units (GPGPUs) enable researchers in a variety of fields to benefit from the computational power of all the cores available inside graphics cards. Understand the Benefits of Using GPUs for Many Scientific Applications **Designing Scientific Applications on GPUs** shows you how to use GPUs for applications in diverse scientific fields, from physics and mathematics to computer science. The book explains the methods necessary for designing or porting your scientific application on GPUs. It will improve your knowledge about image processing, numerical applications, methodology to design efficient applications, optimization methods, and much more. **Everything You Need to Design/Port Your Scientific Application on GPUs** The first part of the book introduces the GPUs and Nvidia's CUDA programming model, currently the most widespread environment for designing GPU applications. The

second part focuses on significant image processing applications on GPUs. The third part presents general methodologies for software development on GPUs and the fourth part describes the use of GPUs for addressing several optimization problems. The fifth part covers many numerical applications, including obstacle problems, fluid simulation, and atomic physics models. The last part illustrates agent-based simulations, pseudorandom number generation, and the solution of large sparse linear systems for integer factorization. Some of the codes presented in the book are available online.

Advancements in microprocessor architecture, interconnection technology, and software development have fueled rapid growth in parallel and distributed computing. However, this development is only of practical benefit if it is accompanied by progress in the design, analysis and programming of parallel algorithms. This concise textbook provides, in one place, three mainstream parallelization approaches, Open MPP, MPI and OpenCL, for multicore computers, interconnected computers and graphical processing units. An overview of practical parallel computing and principles will enable the reader to design efficient parallel programs for solving various computational problems on state-of-the-art personal computers and computing clusters. Topics covered range from parallel algorithms, programming tools, OpenMP, MPI and OpenCL, followed by experimental measurements of parallel programs' run-times, and by engineering analysis of obtained results for improved parallel execution performances. Many examples and exercises support the exposition.

CUDA is a computing architecture designed to facilitate the development of parallel programs. In conjunction with a comprehensive software platform, the CUDA Architecture enables programmers to draw on the immense power of graphics processing units (GPUs) when building high-performance applications. GPUs, of course, have long been available for demanding graphics and game applications. CUDA now brings this valuable resource to programmers working on applications in other domains, including science, engineering, and finance. No knowledge of graphics programming is required—just the ability to program in a modestly extended version of C. *CUDA by Example*, written by two senior members of the CUDA software platform team, shows programmers how to employ this new technology. The authors introduce each area of CUDA development through working examples. After a concise introduction to the CUDA platform and architecture, as well as a quick-start guide to CUDA C, the book details the techniques and trade-offs associated with each key CUDA feature. You'll discover when to use each CUDA C extension and how to write CUDA software that delivers truly outstanding performance. Major topics covered include Parallel programming Thread cooperation Constant memory and events Texture memory Graphics interoperability Atomics Streams CUDA C on multiple GPUs Advanced atomics Additional CUDA resources All the CUDA software tools you'll need are freely available for download from NVIDIA. <http://developer.nvidia.com/object/cuda-by-example.html>

**Recent Developments in Feature Extraction and Selection Algorithms for Data Science**

**Theory and Real World Applications**

**A Self-Teaching Introduction**

**Professional CUDA C Programming**

**Mastering DPC++ for Programming of Heterogeneous Systems using C++ and SYCL**

**Hands-On GPU Programming with Python and CUDA**

If you need to learn CUDA but don't have experience with parallel computing, *CUDA Programming: A Developer's Introduction* offers a detailed guide to CUDA with a grounding in parallel fundamentals. It starts by introducing CUDA and bringing you up to speed on GPU parallelism and hardware, then delving into CUDA installation. Chapters on core concepts including threads, blocks, grids, and memory focus on both parallel and CUDA-specific issues. Later, the book demonstrates CUDA in practice for optimizing applications, adjusting to new hardware, and solving common problems. Comprehensive introduction to parallel programming with CUDA, for readers new to both Detailed instructions help readers optimize the CUDA software development kit Practical techniques illustrate working with memory, threads, algorithms, resources, and more Covers CUDA on multiple hardware platforms: Mac, Linux and Windows with several NVIDIA chipsets Each chapter includes exercises to test reader knowledge Learn how to accelerate C++ programs using data parallelism. This open access book enables C++ programmers to be at the forefront of this exciting and important new development that is helping to push computing to new levels. It is full of practical advice, detailed explanations, and code examples to illustrate key topics. Data parallelism in C++ enables access to parallel resources in a modern heterogeneous system, freeing you from being locked into any particular computing device. Now a single C++ application can use any combination of devices—including GPUs, CPUs, FPGAs and AI ASICs—that are suitable to the problems at hand. This book begins by introducing data parallelism and foundational topics for effective use of the SYCL standard from the Khronos Group and Data Parallel C++ (DPC++), the open source compiler used in this book. Later chapters cover advanced topics including error handling, hardware-specific programming, communication and synchronization, and memory model considerations. Data Parallel C++ provides you with everything needed to use SYCL for programming heterogeneous systems. What You'll Learn Accelerate C++ programs using data-parallel programming Target multiple device types (e.g. CPU, GPU, FPGA) Use SYCL and SYCL compilers Connect with computing 's heterogeneous future via Intel 's oneAPI initiative Who This Book Is For Those new data-parallel programming and computer programmers interested in data-parallel programming using C++.

An In-Depth, Practical Guide to GPGPU Programming Using Direct3D 11 GPGPU Programming for Games and Science demonstrates how to achieve the following requirements to tackle practical problems in computer science and software engineering: Robustness Accuracy Speed Quality source code that is easily maintained, reusable, and readable The book primarily addresses programming on a graphics processing unit (GPU) while covering some material also relevant to programming on a central processing unit (CPU). It discusses many concepts of general purpose GPU (GPGPU) programming and presents

practical examples in game programming and scientific programming. The author first describes numerical issues that arise when computing with floating-point arithmetic, including making trade-offs among robustness, accuracy, and speed. He then shows how single instruction multiple data (SIMD) extensions work on CPUs since GPUs also use SIMD. The core of the book focuses on the GPU from the perspective of Direct3D 11 (D3D11) and the High Level Shading Language (HLSL). This chapter covers drawing 3D objects; vertex, geometry, pixel, and compute shaders; input and output resources for shaders; copying data between CPU and GPU; configuring two or more GPUs to act as one; and IEEE floating-point support on a GPU. The book goes on to explore practical matters of programming a GPU, including code sharing among applications and performing basic tasks on the GPU. Focusing on mathematics, it next discusses vector and matrix algebra, rotations and quaternions, and coordinate systems. The final chapter gives several sample GPGPU applications on relatively advanced topics. Web Resource Available on a supporting website, the author's fully featured Geometric Tools Engine for computing and graphics saves you from having to write a large amount of infrastructure code necessary for even the simplest of applications involving shader programming. The engine provides robust and accurate source code with SIMD when appropriate and GPU versions of algorithms when possible.

The success of Angry Birds, Peggle, and Fruit Ninja has proven that fun and immersive game experiences can be created in two dimensions. Furthermore, 2D graphics enable developers to quickly prototype ideas and mechanics using fewer resources than 3D. 2D Graphics Programming for Games provides an in-depth single source on creating 2D graphics that c

A Developer's Guide to Parallel Computing with GPUs

Modern Java in Action

Programming Models for Parallel Computing

Professional Cuda C Programming

Multicore and GPU Programming

Parallel Computing for Data Science

This book is the result of several years of research trying to better characterize parallel genetic algorithms (pGAs) as a powerful tool for optimization, search, and learning. Readers can learn how to solve complex tasks by reducing their high computational times. Dealing with two scientific fields (parallelism and GAs) is always difficult, and the book seeks at gracefully introducing from basic concepts to advanced topics. The presentation is structured in three parts. The first one is targeted to the algorithms themselves, discussing their components, the physical parallelism, and best practices in using and evaluating them. A second part deals with the theory for pGAs, with an eye on theory-to-practice issues. A final third part offers a very wide study of pGAs as practical problem solvers, addressing domains such as natural language processing, circuits design, scheduling, and genomics. This volume will be helpful both for researchers and practitioners. The first part shows pGAs to either beginners and mature researchers looking for a unified view of the two fields: GAs and parallelism. The second part partially solves (and also opens) new investigation lines in theory of pGAs. The third part can be accessed independently for readers interested in applications. The result is an excellent source of information on the state of the art and future developments in parallel GAs.

A variety of programming models relevant to scientists explained, with an emphasis on how programming constructs map to parts of the computer. What makes computer programs fast or slow? To answer this question, we have to get behind the abstractions of programming languages and look at how a computer really works. This book examines and explains a variety of scientific programming models (programming models relevant to scientists) with an emphasis on how programming constructs map to different parts of the computer's architecture. Two themes emerge: program speed and program modularity. Throughout this book, the premise is to "get under the hood," and the discussion is tied to specific programs. The book digs into linkers, compilers, operating systems, and computer architecture to understand how the different parts of the computer interact with programs. It begins with a review of C/C++ and explanations of how libraries, linkers, and Makefiles work. Programming models covered include Pthreads, OpenMP, MPI, TCP/IP, and CUDA. The emphasis on how computers work leads the reader into computer architecture and occasionally into the operating system kernel. The operating system studied is Linux, the preferred platform for scientific computing. Linux is also open source, which allows users to peer into its inner workings. A brief appendix provides a useful table of machines used to time programs. The book's website (<https://github.com/divakarvi/bk-spca>) has all the programs described in the book as well as a link to the html text.

Summary OpenCL in Action is a thorough, hands-on presentation of OpenCL, with an eye toward showing developers how to build high-performance applications of their own. It begins by presenting the core concepts behind OpenCL, including vector computing, parallel programming, and multi-threaded operations, and then guides you step-by-step from simple data structures to complex functions. About the Technology Whatever system you have, it probably has more raw processing power than you're using. OpenCL is a high-performance programming language that maximizes computational power by executing on CPUs, graphics processors, and other number-crunching devices. It's perfect for speed-sensitive tasks like vector computing, matrix operations, and graphics acceleration. About this Book OpenCL in Action blends the theory of parallel computing with the practical reality of building high-performance applications using OpenCL. It first guides you through the fundamental data structures in an intuitive manner. Then, it explains techniques for high-speed sorting, image processing, matrix operations, and fast Fourier transform. The book concludes with a deep look at the all-important subject of graphics acceleration. Numerous challenging examples give you different ways to experiment with working code. A background in C or C++ is helpful, but no prior exposure to OpenCL is needed.

Purchase of the print book comes with an offer of a free PDF, ePub, and Kindle eBook from Manning. Also available is all code from the book. What's Inside Learn OpenCL step by step  
 Table of Contents PART 1 FOUNDATIONS OF OPENCL PROGRAMMING  
 Introducing OpenCL Host programming: fundamental data structures  
 Host programming: data transfer and partitioning  
 Kernel programming: data types and device memory  
 Kernel programming: operators and functions  
 Image processing  
 Events, profiling, and synchronization  
 Development with C++  
 Development with Java and Python  
 General coding principles  
 PART 2 CODING PRACTICAL ALGORITHMS IN OPENCL  
 Reduction and sorting  
 Matrices and QR decomposition  
 Sparse matrices  
 Signal processing and the fast Fourier transform  
 PART 3 ACCELERATING OPENGL WITH OPENCL  
 Combining OpenCL and OpenGL  
 Textures and renderbuffers  
 Build real-world applications with Python 2.7, CUDA 9, and CUDA 10. We suggest the use of Python 2.7 over Python 3.x, since Python 2.7 has stable support across all the libraries we use in this book.  
 Key Features  
 Expand your background in GPU programming—PyCUDA, scikit-cuda, and Nsight  
 Effectively use CUDA libraries such as cuBLAS, cuFFT, and cuSolver  
 Apply GPU programming to modern data science applications  
 Book Description  
 Hands-On GPU Programming with Python and CUDA hits the ground running: you'll start by learning how to apply Amdahl's Law, use a code profiler to identify bottlenecks in your Python code, and set up an appropriate GPU programming environment. You'll then see how to "query" the GPU's features and copy arrays of data to and from the GPU's own memory. As you make your way through the book, you'll launch code directly onto the GPU and write full blown GPU kernels and device functions in CUDA C. You'll get to grips with profiling GPU code effectively and fully test and debug your code using Nsight IDE. Next, you'll explore some of the more well-known NVIDIA libraries, such as cuFFT and cuBLAS. With a solid background in place, you will now apply your new-found knowledge to develop your very own GPU-based deep neural network from scratch. You'll then explore advanced topics, such as warp shuffling, dynamic parallelism, and PTX assembly. In the final chapter, you'll see some topics and applications related to GPU programming that you may wish to pursue, including AI, graphics, and blockchain. By the end of this book, you will be able to apply GPU programming to problems related to data science and high-performance computing. What you will learn  
 Launch GPU code directly from Python  
 Write effective and efficient GPU kernels and device functions  
 Use libraries such as cuFFT, cuBLAS, and cuSolver  
 Debug and profile your code with Nsight and Visual Profiler  
 Apply GPU programming to datascience problems  
 Build a GPU-based deep neuralnetwork from scratch  
 Explore advanced GPU hardware features, such as warp shuffling  
 Who this book is for  
 Hands-On GPU Programming with Python and CUDA is for developers and data scientists who want to learn the basics of effective GPU programming to improve performance using Python code. You should have an understanding of first-year college or university-level engineering mathematics and physics, and have some experience with Python as well as in any C-based programming language such as C, C++, Go, or Java.

Modern Data Mining Algorithms in C++ and CUDA C

The Developer's Guide

CUDA for Engineers

C Programming

An Introduction to High-Performance Parallel Computing

The Future of Computing Performance

Still more useful techniques, tips, and tricks for harnessing the power of the new generation of powerful GPUs.

Heterogeneous Computing with OpenCL 2.0 teaches OpenCL and parallel programming for complex systems that may include a variety of device architectures: multi-core CPUs, GPUs, and fully-integrated Accelerated Processing Units (APUs). This fully-revised edition includes the latest enhancements in OpenCL 2.0 including:

- Shared virtual memory to increase programming flexibility and reduce data transfers that consume resources
- Dynamic parallelism which reduces processor load and avoids bottlenecks
- Improved imaging support and integration with OpenGL

Designed to work on multiple platforms, OpenCL will help you more effectively program for a heterogeneous future. Written by leaders in the parallel computing and OpenCL communities, this book explores memory spaces, optimization techniques, extensions, debugging and profiling. Multiple case studies and examples illustrate high-performance algorithms, distributing work across heterogeneous systems, embedded domain-specific languages, and will give you hands-on OpenCL experience to address a range of fundamental parallel algorithms. Updated content to cover the latest developments in OpenCL 2.0, including improvements in memory handling, parallelism, and imaging support  
 Explanations of principles and strategies to learn parallel programming with OpenCL, from understanding the abstraction models to thoroughly testing and debugging complete applications  
 Example code covering image analytics, web plugins, particle simulations, video editing, performance optimization, and more

Discover the essential building blocks of the most common forms of deep belief networks. At each step this book provides intuitive motivation, a summary of the most important equations relevant to the topic, and concludes with highly commented code for threaded computation on modern CPUs as well as massive parallel processing on computers with CUDA-capable video display cards. The first of three in a series on C++ and CUDA C deep learning and belief nets, Deep Belief Nets in C++ and CUDA C: Volume 1 shows you how the structure of these elegant models is much closer to that of human brains than traditional neural networks; they have a thought process that is capable of learning abstract concepts built from simpler primitives. As such, you'll see that a typical deep belief net can learn to recognize complex patterns by optimizing millions of parameters, yet this model can still be resistant to overfitting. All the routines and algorithms presented in the book are available in the code download, which also contains some libraries of related routines. What You Will Learn  
 Employ deep learning using C++ and CUDA C  
 Work with supervised feedforward networks  
 Implement restricted Boltzmann machines  
 Use generative samplings  
 Discover why these are important  
 Who This Book Is For  
 Those who have at least a basic knowledge of neural networks and some prior programming experience, although some C++ and CUDA C is recommended.

Parallel Computing for Data Science: With Examples in R, C++ and CUDA is one of the first parallel computing books to

concentrate exclusively on parallel data structures, algorithms, software tools, and applications in data science. It includes examples not only from the classic "n observations, p variables" matrix format but also from time series,

An Introduction to Parallel Programming

OpenCL Programming Guide

Explore high-performance parallel computing with CUDA

Effective techniques for processing complex image data in real time using GPUs

Introduction to Parallel Computing

2D Graphics Programming for Games

Multicore and GPU Programming offers broad coverage of the key parallel computing skillsets: multicore CPU programming and "massively parallel" computing. Using threads, OpenMP, MPI, and CUDA, it teaches the design and development of software that take advantage of today's computing platforms incorporating CPU and GPU hardware and explains how to transition from sequential programming to a parallel computing paradigm. Presenting material refined over more than a decade of teaching parallel computing, Gerassimos Barlas minimizes the challenge with multiple examples, extensive case studies, and full source code. Using this book, you can develop programs that run over distributed memory machines using MPI, create multi-threaded applications with either libraries or write optimized applications that balance the workload between available computing resources, and profile and debug programs on multicore machines. Comprehensive coverage of all major multicore programming tools, including threads, OpenMP, MPI, and C. Demonstrates parallel programming design patterns and examples of how different tools and paradigms can be integrated for performance. Particular focus on the emerging area of divisible load theory and its impact on load balancing and distributed systems. Source code, examples, and instructor support materials on the book's companion website

Explore different GPU programming methods using libraries and directives, such as OpenACC, with extension to languages such as C and Python. Key Features: Learn parallel programming principles and practices and performance analysis in GPU computing. Get distributed multi-GPU programming and other approaches to GPU programming. Understand how GPU acceleration in deep learning can improve their performance. Book Description: Compute Unified Device Architecture (CUDA) is NVIDIA's GPU computing platform and application programming interface. It's designed to work with programming languages such as C, C++, and Python. With CUDA, you can leverage a GPU's parallel computing power for a range of high-performance computing applications in the fields of science, healthcare, and deep learning. Learn CUDA Programming will help you learn GPU parallel programming and understand its modern applications. In this book, you'll discover CUDA programming approaches for modern GPU architectures. You'll not only be guided through GPU features and APIs, you'll also learn how to analyze performance with sample parallel programming algorithms. This book will help you optimize the performance of your apps by giving insights into CUDA programming platforms with various libraries, compiler directives (OpenACC), and other languages. As you progress, you'll learn how additional computing power can be generated using multiple GPUs in a box or across multiple boxes. Finally, you'll explore how CUDA accelerates deep learning algorithms, including convolutional neural networks (CNNs) and recurrent neural networks (RNNs). By the end of this CUDA book, you'll be equipped with the skills you need to integrate the power of GPUs into your applications. What you will learn: Understand general GPU operations and programming patterns in CUDA. Uncover the differences between GPU programming and CPU programming. Analyze GPU application performance and implement optimization strategies. Explore GPU programming, profiling, and debugging tools. Grasp parallel programming algorithms and how to implement them. Scale GPU-accelerated applications with multi-GPU and multi-nodes. Delve into GPU programming platforms with accelerated libraries, Python, and OpenACC. Gain insights into deep learning accelerators in CNNs and RNNs using GPUs. Who this book is for: This beginner-level book is for programmers who want to delve into parallel computing, become part of the high-performance computing community and build their own applications. Basic C and C++ programming experience is assumed. For deep learning enthusiasts, this book covers Python, TensorFlow, and practical examples on performance estimation.

Using the new OpenCL (Open Computing Language) standard, you can write applications that access all available programming resources: CPUs, GPUs, and other processors such as DSPs and the Cell/B.E. processor. Already implemented by Apple, AMD, Intel, IBM, and other leaders, OpenCL has outstanding potential for PCs, servers, handheld/embedded devices, high performance computing, and supercomputing systems. This is the first comprehensive, authoritative, and practical guide to OpenCL 1.1 specifically for working developers and system architects. Written by five leading OpenCL authorities, OpenCL Programming Guide covers the entire specification. It reviews the specification and shows how OpenCL can express a wide range of parallel algorithms, and offers complete reference material on both the API and the C programming language. Through complete case studies and downloadable code examples, the authors show how to write complex programs that decompose workloads across many different devices. They also present all the essentials of OpenCL software development, optimization, including probing and adapting to hardware. Coverage includes: Understanding OpenCL's architecture, concepts, goals, and rationale. Programming with OpenCL C and the runtime API. Using buffers, sub-buffers, images, samplers, and event-driven execution. Synchronizing data with OpenGL and Microsoft's Direct3D. Simplifying development with the C++ Wrapper API. Using OpenCL on embedded devices. Profiles to support devices ranging from cellphones to supercomputer nodes. Case studies dealing with physics simulation; image processing, such as image histograms, edge detection filters, Fast Fourier Transforms, and optical flow; math libraries, such as matrix multiplication and high-performance sparse matrix multiplication; and more. Source code for this book is available at <https://code.google.com/p/opencl-book-samples/>

CUDA for Engineers gives you direct, hands-on engagement with personal, high-performance parallel computing, enabling you to perform computations on a gaming-level PC that would have required a supercomputer just a few years ago. The authors introduce the CUDA C programming clearly and concisely, quickly guiding you from running sample programs to building your own code. Through complete examples you can build, run, and modify, complemented by additional projects that deepen your understanding, the projects are fully developed, with detailed building instructions for all major platforms. Ideal for any scientist, engineer, or student with introductory programming experience, this guide assumes no specialized background in GPU-based or parallel computing. In addition, the authors also present a refresher on C programming for those who need it. Coverage includes: Preparing your computer to run CUDA. Understanding CUDA's parallelism model and C extensions. Transferring data between CPU and GPU. Managing timing, profiling, and debugging. Creating 2D grids. Interoperating with OpenGL to provide real-time user interactivity. Performing basic computations with differential equations. Using stencils to manage related computations across threads. Exploiting CUDA's shared memory to enhance performance. Interacting with 3D data: slicing, volume rendering, and ray casting. Using CUDA libraries. Finding more C resources and code. Realistic example applications include: Visualizing functions in 2D and 3D. Solving differential equations with initial or boundary conditions. Viewing/processing images or image stacks. Computing inner products and centroids. Solving systems of linear equations.

algebraic equations Monte-Carlo computations

GPGPU Programming for Games and Science

C++ AMP

With Examples in R, C++ and CUDA

Deep Belief Nets in C++ and CUDA C: Volume 1

GPU Gems 3

Designing Scientific Applications on GPUs

An Introduction to Parallel Programming, Second Edition presents a tried-and-true tutorial approach that shows students how to develop effective parallel programs with MPI, Pthreads and OpenMP. As the first undergraduate text to directly address compiling and running parallel programs on multi-core and cluster architecture, this second edition carries forward its clear explanations for designing, debugging and evaluating the performance of distributed and shared-memory programs while adding coverage of accelerators via new content on GPU programming and heterogeneous programming. New and improved user-friendly exercises teach students how to compile, run and modify example programs. Takes a tutorial approach, starting with small programming examples and building progressively to more challenging examples Explains how to develop parallel programs using MPI, Pthreads and OpenMP programming models A robust package of online ancillaries for instructors and students includes lecture slides, solutions manual, downloadable source code, and an image bank New to this edition: New chapters on GPU programming and heterogeneous programming New examples and exercises related to parallel algorithms

Discover how CUDA allows OpenCV to handle complex and rapidly growing image data processing in computer and machine vision by accessing the power of GPU Key Features Explore examples to leverage the GPU processing power with OpenCV and CUDA Enhance the performance of algorithms on embedded hardware platforms Discover C++ and Python libraries for GPU acceleration Book Description Computer vision has been revolutionizing a wide range of industries, and OpenCV is the most widely chosen tool for computer vision with its ability to work in multiple programming languages. Nowadays, in computer vision, there is a need to process large images in real time, which is difficult to handle for OpenCV on its own. This is where CUDA comes into the picture, allowing OpenCV to leverage powerful NVIDIA GPUs. This book provides a detailed overview of integrating OpenCV with CUDA for practical applications. To start with, you ' ll understand GPU programming with CUDA, an essential aspect for computer vision developers who have never worked with GPUs. You ' ll then move on to exploring OpenCV acceleration with GPUs and CUDA by walking through some practical examples. Once you have got to grips with the core concepts, you ' ll familiarize yourself with deploying OpenCV applications on NVIDIA Jetson TX1, which is popular for computer vision and deep learning applications. The last chapters of the book explain PyCUDA, a Python library that leverages the power of CUDA and GPUs for accelerations and can be used by computer vision developers who use OpenCV with Python. By the end of this book, you ' ll have enhanced computer vision applications with the help of this book's hands-on approach. What you will learn Understand how to access GPU device properties and capabilities from CUDA programs Learn how to accelerate searching and sorting algorithms Detect shapes such as lines and circles in images Explore object tracking and detection with algorithms Process videos using different video analysis techniques in Jetson TX1 Access GPU device properties from the PyCUDA program Understand how kernel execution works Who this book is for This book is a go-to guide for you if you are a developer working with OpenCV and want to learn how to process more complex image data by exploiting GPU processing. A thorough understanding of computer vision concepts and programming languages such as C++ or Python is expected.

The end of dramatic exponential growth in single-processor performance marks the end of the dominance of the single microprocessor in computing. The era of sequential computing must give way to a new era in which parallelism is at the forefront. Although important scientific and engineering challenges lie ahead, this is an opportune time for innovation in programming systems and computing architectures. We have already begun to see diversity in computer designs to optimize for such considerations as power and throughput. The next generation of discoveries is likely to require advances at both the hardware and software levels of computing systems. There is no guarantee that we can make parallel computing as common and easy to use as yesterday's sequential single-processor computer systems, but unless we aggressively pursue efforts suggested by the recommendations in this book, it will be "game over" for growth in computing performance. If parallel programming and related software efforts fail to become widespread, the development of exciting new applications that drive the computer industry will stall; if such innovation stalls, many other parts of the economy will follow suit. The Future of Computing Performance describes the factors that have led to the future limitations on growth for single processors that are based on complementary metal oxide semiconductor (CMOS) technology. It explores challenges inherent in parallel computing and architecture, including ever-increasing power consumption and the escalated requirements for heat dissipation. The book delineates a research, practice, and education agenda to help overcome these challenges. The Future of Computing Performance will guide researchers, manufacturers, and information technology professionals in the right direction for sustainable growth in computer performance, so that we may all enjoy the next level of benefits to society.

Algorithmic puzzles are puzzles involving well-defined procedures for solving problems. This book will provide an enjoyable and accessible introduction to algorithmic puzzles that will develop the reader's algorithmic thinking. The first part of this book is a tutorial on algorithm design strategies and analysis techniques. Algorithm design strategies — exhaustive search, backtracking, divide-and-conquer and a few others — are general approaches to designing step-by-step instructions for solving problems. Analysis techniques are methods for investigating such procedures to answer questions about the ultimate result of the procedure or how many steps are executed before the procedure stops. The

discussion is an elementary level, with puzzle examples, and requires neither programming nor mathematics beyond a secondary school level. Thus, the tutorial provides a gentle and entertaining introduction to main ideas in high-level algorithmic problem solving. The second and main part of the book contains 150 puzzles, from centuries-old classics to newcomers often asked during job interviews at computing, engineering, and financial companies. The puzzles are divided into three groups by their difficulty levels. The first fifty puzzles in the Easier Puzzles section require only middle school mathematics. The sixty puzzle of average difficulty and forty harder puzzles require just high school mathematics plus a few topics such as binary numbers and simple recurrences, which are reviewed in the tutorial. All the puzzles are provided with hints, detailed solutions, and brief comments. The comments deal with the puzzle origins and design or analysis techniques used in the solution. The book should be of interest to puzzle lovers, students and teachers of algorithm courses, and persons expecting to be given puzzles during job interviews.

Parallel Genetic Algorithms

How to accelerate graphics and computations

CUDA Programming

Lambdas, streams, functional and reactive programming

A Hands-on Approach

Programming Massively Parallel Processors

*Break into the powerful world of parallel GPU programming with this down-to-earth, practical guide. Designed for professionals across multiple industrial sectors, Professional CUDA C Programming presents CUDA -- a parallel computing platform and programming model designed to ease the development of GPU programming -- fundamentals in an easy-to-follow format, and teaches readers how to think in parallel and implement parallel algorithms on GPUs. Each chapter covers a specific topic, and includes workable examples that demonstrate the development process, allowing readers to explore both the "hard" and "soft" aspects of GPU programming. Computing architectures are experiencing a fundamental shift toward scalable parallel computing motivated by application requirements in industry and science. This book demonstrates the challenges of efficiently utilizing compute resources at peak performance, presents modern techniques for tackling these challenges, while increasing accessibility for professionals who are not necessarily parallel programming experts. The CUDA programming model and tools empower developers to write high-performance applications on a scalable, parallel computing platform: the GPU. However, CUDA itself can be difficult to learn without extensive programming experience. Recognized CUDA authorities John Cheng, Max Grossman, and Ty McKercher guide readers through essential GPU programming skills and best practices in Professional CUDA C Programming, including: CUDA Programming Model GPU Execution Model GPU Memory model Streams, Event and Concurrency Multi-GPU Programming CUDA Domain-Specific Libraries Profiling and Performance Tuning. The book makes complex CUDA concepts easy to understand for anyone with knowledge of basic software development with exercises designed to be both readable and high-performance. For the professional seeking entrance to parallel computing and the high-performance computing community, Professional CUDA C Programming is an invaluable resource, with the most current information available on the market. GPU Parallel Program Development using CUDA teaches GPU programming by showing the differences among different families of GPUs. This approach prepares the reader for the next generation and future generations of GPUs. The book emphasizes concepts that will remain relevant for a long time, rather than concepts that are platform-specific. At the same time, the book also provides platform-dependent explanations that are as valuable as generalized GPU concepts. The book consists of three separate parts; it starts by explaining parallelism using CPU multi-threading in Part I. A few simple programs are used to demonstrate the concept of dividing a large task into multiple parallel sub-tasks and mapping them to CPU threads. Multiple ways of parallelizing the same task are analyzed and their pros/cons are studied in terms of both core and memory operation. Part II of the book introduces GPU massive parallelism. The same programs are parallelized on multiple Nvidia GPU platforms and the same performance analysis is repeated. Because the core and memory structures of CPUs and GPUs are different, the results differ in interesting ways. The end goal is to make programmers aware of all the good ideas, as well as the bad ideas, so readers can apply the good ideas and avoid the bad ideas in their own programs. Part III of the book provides pointer for readers who want to expand their horizons. It provides a brief introduction to popular CUDA libraries (such as cuBLAS, cuFFT, NPP, and Thrust), the OpenCL programming language, an overview of GPU programming using other programming languages and API libraries (such as Python, OpenCV, OpenGL, and Apple's Swift and Metal,) and the deep learning library cuDNN.*

*In view of the growing presence and popularity of multicore and manycore processors, accelerators, and coprocessors, as well as clusters using such computing devices, the development of efficient parallel applications has become a key challenge to be able to exploit the performance of such systems. This book covers the scope of parallel programming for modern high performance computing systems. It first discusses selected and popular state-of-the-art computing devices and systems available today, These include multicore CPUs, manycore (co)processors, such as Intel Xeon Phi, accelerators, such as GPUs, and clusters, as well as programming models supported on these platforms. It next introduces parallelization through important programming paradigms, such as master-slave, geometric Single Program Multiple Data (SPMD) and divide-and-conquer. The practical and useful elements of the most popular and important APIs for programming parallel HPC systems are discussed, including MPI, OpenMP, Pthreads, CUDA, OpenCL, and OpenACC. It also demonstrates, through selected code listings, how selected APIs can be used to implement important programming paradigms. Furthermore, it shows how the codes can be compiled and executed in a Linux environment. The book also presents hybrid codes that integrate selected APIs for potentially*

*multi-level parallelization and utilization of heterogeneous resources, and it shows how to use modern elements of these APIs. Selected optimization techniques are also included, such as overlapping communication and computations implemented using various APIs. Features: Discusses the popular and currently available computing devices and cluster systems Includes typical paradigms used in parallel programs Explores popular APIs for programming parallel applications Provides code templates that can be used for implementation of paradigms Provides hybrid code examples allowing multi-level parallelization Covers the optimization of parallel programs*

*Scientific Programming and Computer Architecture*  
*CUDA Application Design and Development*  
*OpenCL in Action*  
*USB Embedded Hosts*  
*Parallel Programming for Modern High Performance Computing Systems*  
*Restricted Boltzmann Machines and Supervised Feedforward Networks*