

Riyadh Baghdadi



Webpage: people.csail.mit.edu/baghdadi/
LinkedIn: www.linkedin.com/in/mrbaghdadi/
Google Scholar: [Link](#)

New York University Abu Dhabi (NYUAD)
A2-185, Saadiyat Island, UAE
+971565262837, baghdadi@nyu.edu

Research Areas

- Intersection of **applied machine learning** and **compilers**:
 - **Machine learning** for **compilers** (e.g., automatic code optimization, and deep learning based heuristics).
 - **Compilers** for **machine learning** (compilers for deep learning hardware accelerators and for optimizing ML).
 - **Compilers** for **high performance computing, tensor algebra, image processing, ...**
-

Education and Postdoctoral Training

Postdoc **Massachusetts Institute of Technology, USA**
Postdoctoral Associate (2015 - 2020).
Advisor: Saman Amarasinghe.

Ph.D. **Sorbonne University, France (ranked 1st in France).**
Ph.D., in Computer Science (2011 - 2015).
Affiliated to INRIA and ENS.
Advisor: Albert Cohen (Google Brain).

M.S. **Sorbonne University, France (ranked 1st in France).**
M.S., in Computer Science (2010 - 2011).
Advisor: Albert Cohen (Google Brain).

Eng. **Ecole Superieure d'Informatique, Algeria (ranked 1st in Algeria).**
Engineering Degree in Computer Science (2005 to 2010).

Professional Experience

NYU - Assistant Professor at New York University Abu Dhabi (2021 - Present).

– **Global Network Assistant Professor at New York University (2021 - Present).**

MIT - Research Affiliate at Massachusetts Institute of Technology (2021 - Present).

- Leading a research team working around the intersection of compilers, deep learning and high performance computing.
- Teaching CSO (Computer Systems Organization - CS-UH-2010). This course provides an introduction to systems. Topics taught include data representation (bits, bytes, floats, strings, etc.), low level machine programming (C and assembly), code optimization, memory hierarchy, virtual memory, webservices and cloud computing.

MIT - Postdoctoral Associate (2015 - 2020).

- Led the team that developed Tiramisu (<http://tiramisu-compiler.org/>), a polyhedral compiler for deep learning, tensor algebra and image processing. It is the first polyhedral compiler to generate code that matches or outperforms highly optimized industrial linear algebra and deep learning libraries (in particular, the most challenging single-kernels from the Intel MKL and MKL-DNN libraries). Compared to state-of-the-art DNN compilers, Tiramisu has the unique ability to parallelize multi-layer RNNs and to exploit DNN weight sparsity. It is open source and has a growing community of contributors from more than 10 industrial and academic institutions.
- Led a team that developed an accurate cost model for automatic code optimization. The cost model is built using deep learning and used to search the space of compiler optimizations and choose the best combination of optimizations. It is the first deep learning based cost model for full programs that does not require heavy feature engineering.
- Collaborated with researchers from UC Berkeley and Facebook to develop an automatic code optimization technique for the Halide compiler. This technique relies on deep learning to create a cost model.
- Member of multiple teams/projects: GraphIt (a DSL compiler for graph analytics), Seq (a python compiler for genomics' applications) and FROST (a DSL backend for FPGA).

Sorbonne University (Paris) - Graduate Student (2011 - 2015).

- Led the team that designed the PENCIL language, a subset of C crafted to allow compilers to generate efficient code for hardware accelerators. This project was a part of the CARP European project and led to many collaborations with industrials and academic institutions including ARM, and Imperial College.
- Developed a technique that relaxes the tiling legality check to allow compilers to apply loop tiling even if the code has false dependences.
- Developed a technique that improves the scalability of automatic optimization in polyhedral compilers (accelerates the Pluto algorithm by $100\times$ on average).
- Member of the team that developed the PENCIL/PPCG polyhedral compiler. Worked mainly on GPU code generation (OpenCL). PENCIL/PPCG is an automatic polyhedral compiler that was recently used to build Tensor Comprehensions, the Facebook deep learning compiler.

NVIDIA - Internship (4 months during 2013).

- Participated in the development of the Nvidia NOVA DSL compiler. A fully automatic DSL for image processing targeting CUDA.

Google - Google Summer of Code (3 months during 2010).

- Designed and implemented a pass to import/export the polyhedral representation of programs from GCC.

Awards

1. **Meta Research Award 2022 (AI4AI Award)**. Award amount \$48k ([Link](#)).
2. **Outstanding Paper Award** in the Fourth Conference on Machine Learning and Systems, 2021.
3. **Bronze medal** in the ACM Student Research Competition (CGO 2015).
4. **Bronze medal** in the ACM Student Research Competition (CGO 2012).
5. **Google summer of code 2010** (\$5000).
6. **Excellence award** for publishing a paper while still a sophomore student (2nd year undergraduate student).
7. **Ranked second nationwide** in the **Microsoft Imagine Cup** competition (2007 and 2008).

Teaching Experience

NYU • Computer Systems Organization (CS-UH-2010). This course provides an introduction to systems and shows students the basics of how a computer works. Topics taught include data representation (bits, bytes, floats, strings, etc.), low level machine programming (C and assembly), code optimization, memory hierarchy, virtual memory, and memory management.

MIT • MIT Kaufman Teaching Certificate Program. Trained and certified for teaching by the MIT KTCP Program.

- **Introduction to Deep Learning.** Teaching assistant in a class of 100 computer science students.
- **Programming High-performance Systems.** Gave a guest lecture about the Tiramisu compiler.

Sorbonne University

- **Introduction to Programming (C Language).** Teaching assistant in a class of 40 undergraduate computer science students.
- **Imperative Programming in Fortran.** Teaching assistant in a class of 40 undergraduate mechanical engineering students.
- **GPGPU programming (CUDA).** Teaching assistant in a class of 30 graduate computer science students.

ESI • Programming High-performance Systems. Principal instructor in a class of 60 graduate computer science students. The class included programming shared memory systems (OpenMP), programming distributed memory systems (MPI), programming GPUs (CUDA), and code optimization techniques. Responsibilities included designing and teaching the whole class.

Papers

Published peer-reviewed papers in conference/journal papers in the venues OOPSLA, SIGGRAPH, ACM TACO, IEEE, PACT, CGO, CC, MLSys, ASAP, ICCD and LCTES. The Google Scholar list of publications is [here](#).

Conference and Journal Papers

1. *"A Deep Learning Model for Loop Interchange."*
CC'2022 (ACM SIGPLAN 2023 International Conference on Compiler Construction).
Lina Mezdour, Khadidja Kadem, Massinissa Merouani, Amina Selma Haichour, Saman Amarasinghe, [Riyadh Baghdadi](#).
Acceptance rate: 39%.
2. *"Q-gym: An Equality Saturation Framework for DNN Inference Exploiting Weight Repetition."* [PDF]
PACT'2022 (International Conference on Parallel Architectures and Compilation Techniques).
Cheng Fu, Hanxian Huang, Bram Wasti, Chris Cummins, [Riyadh Baghdadi](#), Kim Hazelwood, Yuandong Tian, Jishen Zhao, Hugh Leather.
Acceptance rate: 42%.
3. *"Caviar: an e-graph based TRS for automatic code optimization."* [PDF]
CC'2022 (Proceedings of the 31st ACM SIGPLAN International Conference on Compiler Construction).
Smail Kourta, Adel Abderahmane Namani, Fatima Benbouzid-Si Tayeb, Kim Hazelwood, Chris Cummins, Hugh Leather, [Riyadh Baghdadi](#)
Acceptance rate: 32%.
4. *"A Deep Learning Based Cost Model For Automatic Code Optimization."* [PDF]
MLSys'2021 (Fourth Conference on Machine Learning and Systems)
Best Paper Award
[Riyadh Baghdadi](#), M. Merouani, M. H. Leghettas, K. Abdous, T. Arbaoui, K. Benatchba, S. Amarasinghe
Acceptance rate: 23.5%.
5. *"Hardware Acceleration of Sparse and Irregular Tensor Computations of ML Models: A Survey and Insights."* [PDF]
Proceedings of the IEEE 2021
S. Dave, [Riyadh Baghdadi](#), T. Nowatzki, S. Avancha, A. Shrivastava, B. Li
Acceptance rate: 30%.

6. *"Learning to Optimize Halide with Tree Search and Random Programs"*. [PDF][WebSite]
SIGGRAPH 2019 (Special Interest Group on Computer Graphics and Interactive Techniques).
 Andrew A., Karima M., Luke A., Tzu-Mao L., Michael G., Riyadh Baghdadi, Steven J., Benoit S., Jonathan R., Fredo D.
 Acceptance rate: 29%.
7. *"Seq: A high-performance language for computational biology"*. [PDF][WebSite]
OOPSLA 2019 (Object-Oriented Programming, Systems, Languages & Applications).
 A. Shajii, I. Numanagic, Riyadh Baghdadi, B. Berger, S. Amarasingh.
 Acceptance rate: 36%.
8. *"Tiramisu: A Polyhedral Compiler for Expressing Fast and Portable Code"*. [PDF] [WebSite].
CGO'19 (Code Generation and Optimization).
Riyadh Baghdadi, J. Ray, M. B. Romdhane, E. D. Sozzo, A. Akkas, Y. Zhang, P. Suriana, S. Kamil, S. Amarasinghe.
 Acceptance rate: 30%.
9. *"GraphIt - A High-Performance DSL for Graph Analytics"*. [PDF][WebSite]
OOPSLA'18 (Object-Oriented Programming, Systems, Languages & Applications).
 Y. Zhang, M. Yang, Riyadh Baghdadi, S. Kamil, J. Shun, S. Amarasinghe.
 Acceptance rate: 29%.
10. *"A Unified Backend for Targeting FPGAs from DSLs"*. [PDF]
ASAP'18 (2018 IEEE 29th International Conference on Application-specific Systems, Architectures and Processors).
 E. Sozzo, Riyadh Baghdadi, S. Amarasinghe, and M. Santambrogio.
 Acceptance rate: 30%.
11. *"A Common Backend for Hardware Acceleration on FPGA"*. [PDF]
ICCD'17 (37th IEEE International Conference on Computer Design).
 E. D. Sozzo, Riyadh Baghdadi, S. Amarasinghe, M. D. Santambrogio.
 Acceptance rate: 29%.
12. *"PENCIL: a Platform-Neutral Compute Intermediate Language for Accelerator Programming"*. [PDF][WebSite]
PACT'15 (International Conference on Parallel Architectures and Compilation Techniques).
Riyadh Baghdadi, U. Beaugnon, A. Cohen, T. Grosser, M. Kruse, C. Reddy, S. Verdoolaege, J. Absar, S. v. Haastregt, A. Kravets,
 A. Lokhmotov, A. Betts, J. Ketema, A. F.~Donaldson, R. David, E. Hajiyev.
 Acceptance rate: 21%.
13. *"VOBLA: A Vehicle for Optimized Basic Linear Algebra"*. [PDF]
LCTES'14 (Languages, Compilers, Tools and Theory of Embedded Systems).
 U. Beaugnon, A. Kravets, S. V. Haastregt, Riyadh Baghdadi, D. Tweed, J. Absar, A. Lokhmotov.
 Acceptance rate: 31%.
14. *"Improved Loop Tiling Based on the Removal of Spurious False Dependences"*. [PDF]
ACM TACO (ACM Transactions on Architecture and Code Optimization).
Riyadh Baghdadi, A. Cohen, S. Verdoolaege, K. Trifunovic.
 Acceptance rate: 30%.

Workshops

1. *"Progress Report: A Deep Learning Guided Exploration of Affine Unimodular Loop Transformations."* [PDF]
IMPACT'22 (12th International Workshop on Polyhedral Compilation Techniques).
 Massinissa Merouani, Khaled Afif Boudaoud, Iheb Nassim Aouadj, Nassim Tchoulak, Fatima Benbouzid-Sitayeb,
 Karima Benatchba, Hugh Leather, Riyadh Baghdadi.
2. *"SALSA: A Domain Specific Architecture for Sequence Alignment"*. [PDF]
27'th RAW (Reconfigurable Architectures Workshop) at IPDPS 2020, May., 2020, New Orleans, Louisiana, USA.
 L. D. Tucciy, Riyadh Baghdadi, S. Amarasinghe, M. D. Santambrogio.
3. *"Tiramisu: A Polyhedral Compiler for Dense and Sparse Deep Learning"*. [PDF]
Systems for ML workshop at NeurIPS 2019, Dec., 2019, Vancouver, Canada.
Riyadh Baghdadi, K. Abdous, A. N. Debbagh, B. F. Zohra, A. Renda, J. E. Frankle, M. Carbin, S. Amarasinghe.
4. *"Scalable Polyhedral Compilation, Syntax vs. Semantics: 1-0 in the First Round"*. [PDF]
IMPACT (10th International Workshop on Polyhedral Compilation Techniques). Jan., 2020. Bologna, Italy.
Riyadh Baghdadi, A. Cohen.

5. *"Language Support For Better Polyhedral Compilation Targeting Accelerators"*.
GRGPL'15 (Journées du Groupement de Recherche Genie de la Programmation et du Logiciel), France.
[Riyadh Baghdadi](#).
6. *"PENCIL: A platform-neutral intermediate language for the parallelizing compilation of DSLs"*.
DSLDI'14 (Domain-Specific Language Design and Implementation), Portland, Oregon, USA.
U. Beaugnon, [Riyadh Baghdadi](#), J. Absar, A. Betts, A. Cohen, A. Donaldson, T. Grosser, S. V. Haastregt, Y. Hu, J. Ketema, A. Kravets, A. Lokhmotov, S. Verdoolaege.
7. *"Pencil: Towards a Platform-Neutral Compute Intermediate Language for DSLs"*.
WOLFHPC'12 (DSLs and High-Level Frameworks for High Performance Computing), in conjunction with OOPSLA.
[Riyadh Baghdadi](#), A. Cohen, S. Guelton, S. Verdoolaege, J. Inoue, T. Grosser, G. Kouveli, A. Kravets, A. Lokhmotov, C. Nugteren, F. Waters, A. F.~Donaldson. [PDF]
8. *"The Potential of Synergistic Static, Dynamic and Speculative Loop Nest Optimizations for Automatic Parallelization"*.
PESPMA'10 (Parallel Execution of Sequential Programs on Multi-core Architectures), in conjunction with ISCA.
[Riyadh Baghdadi](#), A. Cohen, C. Bastoul, L-N. Pouchet and L. Rauchwerger.

PhD Thesis

1. *Improving Tiling, Reducing Compilation Time, and Extending the Scope of Polyhedral Compilation*, Riyadh Baghdadi.

Reports

1. *"PENCIL Language Specification"*.
Research Report RT-8706, INRIA.
[Riyadh Baghdadi](#), A. Cohen, S. Verdoolaege, T. Grosser, J. Absar, S. v. Haastregt, A. Kravets, A. Lokhmotov, A. F.~Donaldson.
[PDF]

Presentations, Posters and Invited Talks

Presentations, posters and invited talks in the areas of code generation for deep learning, automatic optimization using deep learning, compilers, and code optimization.

Tiramisu - Deep Learning Based Cost Model for Automatic Code Optimization

1. ML for Systems workshop 2022 (Neurips'22), Dec. 2022.
2. Huawei, Feb. 2022.
3. Google Brain, Feb. 2021.

Tiramisu - Accelerating Deep Learning

1. Nvidia, Dec. 2019.
2. MIT, FastCode Seminar, Oct. 2019.
3. Microsoft, Aug. 2019.
4. BigStream, Aug. 2019.
5. Cerebras, Aug. 2019.
6. SRC TECHCON, Sep. 2019 (Austin, TX, USA).
7. Apple Sep. 2018.
8. 2019 MIT Alliances annual meeting, Jun. 2019 (Cambridge, MA, USA).
9. 2018 MIT Alliances annual meeting, Jun. 2018, (Cambridge, MA, USA).

Tiramisu - General

1. *"A Platform for Exploring Machine Learning Based AutoScheduling"*.
Workshop on Optimization, Modeling, Analysis and Space Exploration, Feb. 2019.
[Riyadh Baghdadi](#).
2. *"Accelerating LQCD Calculations Using the Tiramisu Compiler"*.
Advanced Computing Principal Investigator (PI) Meeting, July 2019, Rockville, MD.
[Riyadh Baghdadi](#), M Wagman, A. Pochinsky, S. Amarasinghe, W. Detmold.
3. *"The Tiramisu Polyhedral Compiler for Deep Learning and Dense Computations"*.
MIT PL Offsite meeting, May 2019, Cambridge, MA, USA.
[Riyadh Baghdadi](#)
4. *"Tiramisu: A High-Performance Compiler for Domain-Specific Architectures"*.
2019 ADA Annual Symposium (Center for Application Driven Architectures), April 2019. Ann Arbor, MI, USA.
[Riyadh Baghdadi](#)

PENCIL - General

1. *"PENCIL: a subset of C99 for Accelerator Programming"*.
LAMIH Seminar, University of Valenciennes, Sep 2015, Valenciennes, France.
[Riyadh Baghdadi](#).
2. *"PENCIL: a Platform-Neutral Compute Intermediate Language for DSL Compilers and for Accelerator Programming"*.
MIT Seminar - Massachusetts Institute of Technology, May 2015, Cambridge, Massachusetts.
[Riyadh Baghdadi](#).
3. *"Extending the Scope of Polyhedral Compilation: Progress in Handling Irregular Codes and in Scalability"*.
ACM SRC, CGO 2015.
[Riyadh Baghdadi](#).
4. *"Generating Highly Optimized CUDA and OpenCL from Domain Specific Languages"*.
Google PhD Student Summit, 2014.
[Riyadh Baghdadi](#).
5. *"Language support for polyhedral compilation: evaluation on image processing benchmark"*.
8'th meeting of the french compiler community, 2014.
[Riyadh Baghdadi](#), S. Verdoolaege, U. Beaugnon, A. Cohen, R. David and E. Hajiyev.
6. *"Putting Polyhedral Optimization Techniques to Work in Production Compilers: Progresses in Scalability and Memory Management"*.
ACM SRC, CGO 2012.
[Riyadh Baghdadi](#).
7. *"A relaxed permutability criterion"*.
Dixiemes rencontres de la communaute francaise de compilation, Sep 2015, Banyuls-sur-Mer, France
[Riyadh Baghdadi](#).

Funding and Grant Awards

- Grants I wrote (with PI status) and were accepted:
 - Program Name: Meta Research Award 2022 (AI4AI Award). Accepted Proposal: Goal-oriented search for un-common code optimizations in Halide. Funding Amount: 48k USD.
- Grants I participated in writing (without PI status) and were accepted:
 - Program name: DARPA D3M (Data-Driven Discovery of Models). Accepted proposal: Human Data Interaction (DARPA-16-43-D3M-FP-041).

- Program name: DARPA PAPP (Performant Automation of Parallel Program Assembly). Accepted proposal: High-Performance Productivity and Portability with Domain Specific Languages (DARPA-PA-19-04-02-FP-019).
- Program name: DOE SCIDAC (Scientific Discovery Through Advanced Computing). Accepted proposal: Computing the Properties of Matter with Leadership Computing Resources (DE-SC0018121).
- Program name: MIT Alliances FinTech funding (FinTech@MIT). Accepted proposal: A DSL for Real-time, Low-latency Financial Computations.
- Program name: Advanced Simulation and Computing Predictive Science Academic Alliance Program (PSAAP III). Accepted proposal: CESMIX - Center for the Exascale Simulation of Material Interfaces in Extreme Environments.

Service

Reviews & Program Committee Membership

1. **Journal reviews:** ACM TACO (Transactions on Architecture and Code Optimization), ACM TOPC (Transactions on Parallel Computing), Elsevier PARCO (Parallel Computing), Elsevier JPDC (Journal of Parallel and Distributed Computing), Springer IJPP (International Journal of Parallel Programming), IEEE Access Journal.
2. **Conference reviews:** PACT'19, ECOOP 2023, IPDPS 2023, MLSys'23.
3. **Workshop reviews:** RWDSL'18 (Workshop on Real World Domain Specific Languages), GPGPU'10 (Workshop on General-Purpose GPU), ML for Systems (Workshop on Machine Learning for Systems at NeurIPS 2022).

Workshop Organization

1. 12th International Workshop on Polyhedral Compilation Techniques, June 20th, 2022, Budapest, Hungary, In conjunction with HiPEAC 2022. [Link](#).
2. The 2nd International Workshop on Machine Learning for Software Hardware Co-Design (MLSH'21), September 26'th, 2021, Virtual, In conjunction with PACT21. [Link](#).
3. The 1st International Workshop on Machine Learning for Software Hardware Co-Design (MLSH'20), October 2nd, 2020, Virtual, In conjunction with PACT20. [Link](#).

NYUAD Service

1. Science Divisional Representative on the Global Education and Community-Based Learning Committee (2022-2023).
2. Member of the following search committees: visiting professors (Spring 2023), visiting professors (Spring 2021), lecturers (Spring 2021), open rank professors (Spring 2021).
3. Co-organization of the NYUAD Hackathon for Social Good (Spring 2022 and Spring 2023).

Student Mentoring

- PhD Students.
 - Massinissa Merouani (NYUAD).
 - Yacine Hakimi (ESI).
- Visiting PhD students (interns).
 - Lorenzo Di Tucci (visiting from Politecnico di Milano, 2018-2019). One paper published.
 - Emanuele Del Sozzo (visiting from Politecnico di Milano, 2017-2018). Two papers published.
- Master students (MIT, USA)

- Abdurahmane Akkas (2018-2019). “Efficient Memory and GPU Operations for Tiramisu Compiler”. [PDF]
- Malek Ben Romdhane (2017-2018). “Extending the Capabilities of Tiramisu”. [PDF]
- Master students (ESI, Algeria)
 - A. Boudaoud, N. Aouadj (2021-2022). “Approche Basée sur l’Apprentissage Profond pour l’Estimation des Performances dans les Compilateurs Polyédriques.”. [PDF]
 - N. Hennouni, N. El Hassane (2021-2022). “Utilisation de l’apprentissage par renforcement dans l’exploration de l’espace de recherche pour l’optimisation automatique des compilateurs.”. [PDF]
 - L. Mezdour, K. Kdem (2021-2022). “Optimisation automatique dans le compilateur Tiramisu : Amélioration de la scalabilité par la sélection des paramètres.”. [PDF]
 - N. Tchoulak (2020-2021). “Automatic code optimization with polyhedral transformations and DeepLearning.”. [PDF]
 - A. Tadjer (2020-2021). “Building a Text-to-Code Model”.
 - H. Benyamina (2020-2021). “A comprehensive programming language model.”. [PDF]
 - S. Kourta, A. Namani (2020-2021). “An E-Graph Based Term Rewriting System for Automatic Code Optimization.”. [PDF]
 - M. Merouani, M. H. Leghettas (2019-2020). “Building a Deep Learning Based Cost Model for Code Optimization in Tiramisu”. [PDF]
 - K. Abdous (2019-2020). “Automatic Code Optimization Algorithm for Tiramisu”. [PDF]
 - H. Benmeziane (2019-2020). “Integrating Tiramisu in the Pytorch DNN Framework”. [PDF]
 - A. N. Debbagh (2019-2020). “Optimizing Sparse DNNs using Tiramisu”. [PDF]
 - T. Ihadadene (2018-2019). “Generating Communication Code for Distributed Programs”. [PDF]
 - M. I. Isra, M. Henni (2018-2019). “A Deep Learning Approach for Automatic Code Optimization”. [PDF]
 - G. Amal (2018-2019). “Generation automatique de communication CPU/GPU pour le backend GPU de Tiramisu”. [PDF]
 - G. Taklit, A. Balamane (2018-2019). “Proposition d’un modèle pour l’optimisation automatique de boucles”. [PDF]
 - H. Radja (2018-2019). “Implementing and Optimizing Neural Networks using Tiramisu”. [PDF]
 - I. Manseri (2017-2018). “Design and implementation of an optimizer for programs written in Halide”.

References

Saman Amarasinghe

MIT, EECS
saman@csail.mit.edu
+1 781-888-0816

Albert Cohen

Google AI
albertcohen@google.com
+33 66 88 62 206

William Detmold

MIT, CTP
wdetmold@mit.edu
+1 617-324-6181