

# Michael Luo

☎ (408)-888-8377 | ✉ [michael.luo@berkeley.edu](mailto:michael.luo@berkeley.edu) | 🌐 [www.michaelzhiluo.com](http://www.michaelzhiluo.com) | 📄 [michaelzhiluo](https://michaelzhiluo.github.io) | **in** [michaelzhiluo](https://www.linkedin.com/in/michaelzhiluo)

## EDUCATION

---

### University of California: Berkeley

Berkeley CA

*PhD in Electrical Engineering and Computer Science, Advised by Ion Stoica*

*Aug 2021 - Present*

- Coursework: Deep Unsupervised Learning (CS 294-158), Deep Reinforcement Learning (CS 285), NLP (CS 288), Computer Vision (CS 280), Theoretical Statistics (STAT 210A), Machine Learning (CS 189), Operating Systems (CS 262), Databases (CS 186)

### University of California: Berkeley

Berkeley, CA

*Summa Cum Laude — GPA: 3.98/4, M.S. in EECS + B.S. in EECS + Business*

*Aug 2016 – May 2021*

## PUBLICATIONS

---

### Balsa: Learning a Query Optimizer Without Expert Demonstrations

Zongheng Yang, Wei-lin Chiang, Frank Luan, Gautam Mittal, **Michael Luo**, Ion Stoica

**SIGMOD 2022**

- An end2end query optimizer trained via deep RL that exceeds the query-performance of expert solvers by up to 2.8x.

### MESA: Offline Meta-RL for Safe Adaptation and Fault Tolerance

**Michael Luo**, Ashwin Balakrishna, Brijen Thananjeyan, Suraj Nair, Julian Ibarz, Jie Tan, Chelsea Finn, Ion Stoica, and Ken Goldberg

**NeurIPS Safe Control Workshop 2021**

- Safe RL algorithm that meta-learns from offline datasets to safely adapt to unseen environments.

### Importance Weighted Asynchronous Architectures with Clipped Target Networks

**Michael Luo**, Jiahao Yao, Richard Liaw, Eric Liang, Ion Stoica

**International Conference on Learning Representations (ICLR) 2020**

- Distributed RL algorithm that combines the sample efficiency of PPO and training speed of IMPALA in model-free RL, resulting in 2x increase in training speed

### Connecting Context-specific Adaptation in Humans to Meta-learning

Rachit Dubey\*, Erin Grant\*, **Michael Luo**\*, Karthik Narasimhan, Thomas L. Griffith

Preprint

- Augmenting gradient-based meta-learning algorithms with conditioning networks with contextual cues as input to improve task-adaptation; Attains 1.2x higher post-adaptation performance than context-concatenated baselines

### Recovery RL: Safe Reinforcement Learning with Learned Recovery Zones

Brijen Thananjeyan\*, Ashwin Balakrishna\*, Suraj Nair, **Michael Luo**, Krishnan Srinivasan, Minh Hwang, Joseph Gonzalez, Julian Ibarz, Chelsea Finn, Ken Goldberg

**NeurIPS Robotic Learning Workshop 2020**

- Safe RL method that employs a safety critic and recovery policy to predict the probability the agent is violating a constraint and return the agent from a dangerous state back to a safe state; Attains up to 50x less total constraint violations than existing safe RL methods.

### Distributed Reinforcement Learning is a Dataflow Problem

Eric Liang\*, Zhanghao Wu\*, **Michael Luo**, Sven Mika, Ion Stoica

**NeurIPS 2021**

- A new execution API that recasts RL algorithms as a hybrid-actor framework with parallel iterators, implemented for 20 RL algorithms, resulting in 2x faster training, 1.2x data throughput, and 2-3x reductions in code.

### Accelerating Quadratic Optimization with Reinforcement Learning

Jeffrey Ichnowski, Paras Jain, Bartolomeo Stellato, Goran Banjac, **Michael Luo**, Francesco Borrelli, Joseph E. Gonzalez, Ion Stoica, Ken Goldberg

**NeurIPS 2021**

- An intelligent application of RL that tunes the parameters of existing Quadratic Program (QP) solvers and improving solving times by up to 3x.

## Discovering Autoregressive Orderings with Variational Inference

Xuanlin Li\*, Brandon Trabucco\*, Dong Huk Park, Yang Gao, Michael Luo, Sheng Shen, Trevor Darrell

International Conference on Learning Representations (ICLR) 2021

- We learn autoregressive orderings from scratch in language modeling through a practical end-to-end algorithm that employs variational inference and RL to infer doubly stochastic matrices; Learned orderings significantly beat standard L2R baselines.

## LazyDagger: Reducing Context Switching in Interactive Robot Imitation Learning

Ryan Hoque, Ashwin Balakrishna, Brijen Thananjeyan, Carl Putterman, Michael Luo, Daniel Seita, Daniel Brown, Ken Goldberg

CASE 2021

## AlphaGarden: Learning Seed Placement and Automation Policies For Polyculture Farming

Yahav Avigal, Anna Deza, William Wong, Sebastian Oehme, Mark Presten, Mark Theis, Jackson Chui, Paul Shao, Huang Huang, Atsunobu Kotani, Satvik Sharma, Michael Luo, Stefano Carpin, Joshua Viers, Stavros Vougioukas, and Ken Goldberg

International Conference on Robotics and Automation (ICRA) 2021

---

## WORK EXPERIENCE

### Software Development Engineer, Intern

June 2020 – Aug 2020

*Anyscale, Ray Core/RLlib Team*

*Berkeley, CA*

- Developed model-based and meta-learning RL algorithms, improving sample efficiency by 10x against model-free RL benchmarks

### Teaching Assistant for Machine Learning Course (CS 189)

Aug 2019 – Dec 2019

*University of California: Berkeley*

*Berkeley, CA*

### Software Development Engineer, Intern

June 2019 – Aug 2019

*Amazon A9, Search Advertisement Team*

*Palo Alto, CA*

- Developed a deep contrastive-learning approach to differentiate and quickly classify advertisement images into multi-class categories

### Computer Vision Engineer, Intern

June 2018 – Aug 2018

*Cisco Meraki, Smart Camera Team*

*San Francisco, CA*

- Experimented with various deep and algorithmic object detection and tracking methods to track and identify individuals across Meraki cameras in public areas

---

## PROJECTS

### Sky Computing Project | *Python*

Sept 2021 – Now

- Developed framework for easily running ML workloads on any cloud.
- Created Sky Storage service that abstracts away storage services such as S3 and GCS.
- Created Sky On-prem feature that allows for users to run jobs on their own clusters and autoscale to clouds.

### Ray RLlib Core Contributor | *Python, Pytorch, Tensorflow*

Aug 2018 – Aug 2021

- Created distributed model-free, model-based, and meta-learning RL algorithms on RLlib, including APPO/IMPACT, MAML, MBMPO, Google Dreamer
- Improved and tuned performance of existing algorithms on various, including PPO and SAC
- Developed SOTA distributed PPO implementation, beating previous SOTA by two-fold, achieving 10k reward on HalfCheetah-v2 in 1 hour

### Spotify Recommendation Systems Project | *Python, Tensorflow*

Feb 2018 – May 2018

- Devised two latent embedding vector approaches from Spotify Million Playlist Dataset, one embedding trained via contrastive learning over pairwise song IOUs and the other trained via Skipgram encoding; sequences are then predicted with LSTM

---

## TECHNICAL SKILLS

**Areas:** Deep Reinforcement Learning, Operating Systems, Robotics, Natural Language Processing

**Languages:** Python, Java, C/C++, SQL (Postgres), Javascript, C#, Golang, Bash, HTML/CSS

**Developer Tools:** Git, Docker, TravisCI, Google Cloud Platform, Kubernetes, Vim, Sublime Text

**Libraries:** Tensorflow, Pytorch, Pandas, Numpy, Matplotlib, Seaborn, OpenAI Gym