Abhi Gupta

abhi.gupta@columbia.edu
www.abhigupta.ai

Education

Columbia University, Fu Foundation School of Engineering and Applied Sciences NY, USA
Bachelor of Science in Engineering 2017-2021

Major: Computer Science, Minor: Operations Research GPA 3.96/4.0
Honors Artificial Intelligence Track

Research Interests

My goal is to leverage reinforcement learning to solve the challenges of real life that cannot be simulated. With a focus on the theory of optimization in sequential decision-making, I hope to extend online algorithms to learn directly from real-world interaction and generalize to all situations, including those that cannot be predetermined.

Publications

[1] **Abhi Gupta**, Yunhao Tang, Shipra Agrawal, and Yuri Faenza. "Formulating Branch-and-Bound as a Reinforcement Learning Problem." **In preparation for ICML 2021.**

[2] **Abhi Gupta**, Jingxi Xu, Peter Allen, and Shuran Song. "A Self-Supervised Framework for Online Robot-Camera Calibration." **In preparation for RSS 2021.**

[3] Robert Kwiatkowski, **Abhi Gupta**, Wonjun Son, Boyuan Chen, and Hod Lipson. "Leveraging Uncertainty-Aware Dynamics Models to Refine Actor-Critic Methods." **In preparation for ICML 2021.**

[4] Bohan Wu, Feng Xu, Zhanpeng He, **Abhi Gupta**, and Peter K. Allen. "SQUIRL: Robust and Efficient Learning from Video Demonstration of Long-Horizon Robotic Manipulation Tasks." **IROS 2020.**

[5] Bohan Wu, Iretiayo Akinola, **Abhi Gupta**, Feng Xu, Jake Varley, David Watkins-Valls, and Peter Allen, "Generative Attention Learning: A General Framework for High-Performance Multi-fingered Grasping in Clutter," **Autonomous Robots 2020.**

Fellowships & Scholarships

[1] **QuestBridge Scholar** 09/2017 – 05/2021

As a low-income high school senior who excelled academically, I was selected to be a QuestBridge National College Match Finalist and offered admission to Columbia University with full financial support.

[2] **Columbia Data Science Institute Summer Scholar** 05/2018 – 09/2018

I was the sole college freshman selected to join the Data Science Institute to explore fairness in AI. As part of the summer cohort, I was selected by Professor Peter Allen to join his research lab in robotics.

[3] **SHAD Valley Fellow** 05/2016 – 09/2016

I was the only high school student in my region selected to attend Dalhousie University in Nova Scotia to pursue a summer leadership and entrepreneurship incubator. I led a team of 54 to execute my business

plan and independently developed a prototype. Among 2,600 other fellows across Canada, I was awarded Best Engineering Design for our product.

Honors and Awards

[1] Awarded Dean's List all semesters
[2] Invited to Tau Beta Pi Engineering Honor Society in Junior Year
[3] Invited to Peer-Review for International Conference on Robotics and Automation (ICRA) 2021

Teaching

[1] **Reinforcement Learning: Theory & Practice** 01/2021 – 05/2021

I was invited by Professor Shipra Agrawal to be the course assistant of the graduate level course *Reinforcement Learning: Theory & Practice (IEOR 4575)* next semester.

[2] **Private Tutoring** 09/2018 – 03/2020

I privately tutored undergraduate and masters students in subjects including physics, multivariable calculus, linear algebra, linear programming, probability and statistics, stochastic models, and advanced programming. I also offered support with training deep learning models and configuring cloud services.

[3] **FIRST Robotics Coach** 09/2017 – 03/2020

I led our high school robotics team to the FIRST Robotics Great Lakes Region 1st place win and top 50 finish in the 2015 St. Louis FRC World Championships. As a former team member with college-level experience in robotics, I provide mentorship and feedback in preparation for the competition every year.

Research Experience

[1] **Formulating Branch-and-Bound as a Reinforcement Learning Problem** NY, USA
Advised by Dr. Shipra Agrawal at Columbia IEOR 06/2020 - ongoing

Problem B&B is the state-of-the-art algorithm for solving NP-hard combinatorial optimization problems. Finding a feasible solution with a certificate of optimality requires two kinds of decisions: how to *branch* on the current problem and how to *select* the next problem. Currently, decision rules in B&B rely on suboptimal heuristics that are prohibitively time-consuming and explore excessively large search trees.

Analysis I show the expert decision rule, known as fullstrong branching, performs well because it benefits from the side effects of solving auxiliary linear programs. With side-effects such as pruning and improvements in primal bound disabled, the decision rule is no longer so great. Nevertheless, seeking improvement in dual bound remains a compelling strategy. Furthermore, I conclude global branching decisions reduce to those based on local heuristics when they have no control over the selection rule.

Algorithm I provided the first RL formulation of B&B that can jointly learn a branching and selecting policy. I introduced a Q-network architecture to parametrize a changing action space with graph neural networks and attention mechanisms. I combatted the growing action space faced in B&B by factorizing the policy into a brancher that is conditioned on the decisions of an independently-trained selector.

Publication **Abhi Gupta**, Yunhao Tang, Shipra Agrawal, and Yuri Faenza. "Formulating Branch-and-Bound as a Reinforcement Learning Problem." **In preparation for ICML 2021.**

[2] **A Self-Supervised Framework for Online Robot-Camera Calibration** NY, USA
Advised by Dr. Shuran Song at Columbia CS 05/2020 - ongoing

PROBLEM Camera extrinsics can be computed from associations between points in the world and pixels in the image. The best procedures today are offline and cumbersome, requiring recalibration every time the camera is moved. Pose tracking can be used in place of markers but needs many real-world labels to be accurate. Calibrating a scene on-the-fly should be possible without anything more than what the camera can see.

ALGORITHM Articulated objects, like robots, can be tracked without collecting real-world labels by leveraging their 3D CAD models instead. Because robots in real-life settings will closely resemble the geometry of their models, I pre-train a network directly in simulation to estimate joint states from a depth image. Next, I fine-tune the network with unlabeled, real-world depth images by first treating them as partial point clouds of the robot, then comparing the partial point cloud to the robot's CAD model, and finally backpropagating a loss through my differentiable forward kinematics.

RESULTS The framework requires no more knowledge than the model of the object in the scene. It is able to both predict joint states and supervise its predictions from the same depth image. In addition, it can estimate the values of underactuated joints whose ground-truth is not available. Real world experiments are currently in progress.

PUBLICATION **Abhi Gupta**, Jingxi Xu, Peter Allen, and Shuran Song. "A Self-Supervised Framework for Online Robot-Camera Calibration." **In preparation for RSS 2021.**

[3] **Leveraging Uncertainty-Aware Dynamics Models to Refine Actor-Critic Methods** NY, USA
Advised by Dr. Hod Lipson at Columbia CS 01/2020 - ongoing

PROBLEM Model-based RL improves the sample complexity of its model-free counterparts at the cost of asymptotic performance. Imperfect models of continuous, high-dimensional dynamics introduce an approximate MDP under which even optimal policies are often useless in the real environment. Compounding errors in model prediction, also known as exposure bias, further create a bottleneck for performance of model-based RL algorithms.

ALGORITHM As in Soft Actor-Critic, the model-free training loop has an experience replay buffer filled with real data. We augmented this training loop with an added planning step of Monte Carlo Tree Search (MCTS). I showed fully Bayesian forward models with recurrent architectures are difficult to train, but those with stochastic latent embeddings, as in variational autoencoders, are more robust to compounding error. I introduced a novel learning curriculum to stabilize training over increasing horizons. For the first time, we used model uncertainty to penalize the rewards of states rolled out during MCTS to avoid trusting imagined trajectories that are unlikely to occur in the real environment.

RESULTS On a series of continuous control OpenAI gym tasks with varying difficulty, we consistently improve the performance of model-free RL even with imperfect models of the world.

PUBLICATION Robert Kwiatkowski, **Abhi Gupta**, Wonjun Sun, Boyuan Chen, and Hod Lipson. "Leveraging Uncertainty-Aware Dynamics Models to Refine Actor-Critic Methods." **In preparation for ICML 2021.**

[4] **Curiosity-Driven Exploration for Real Life Reinforcement Learning** NY, USA
Advised by Dr. Hod Lipson at Columbia CS 07/2020 – 12/2020

PROBLEM Extrinsic rewards are often absent or sparse in real-life decision making. In these circumstances, models can be used not only to plan but also to construct intrinsic, task-

agnostic rewards. These rewards encourage the agent to be curious and to quickly adapt to new tasks. It is an open question whether policies pretrained under curiosity in simulation can learn task-specific behavior in the real world.

EXPERIMENT In experiments conducted on continuous control OpenAI gym tasks, I find the model disagreement metric conventionally used as intrinsic reward is vulnerable to exploitation. We are evaluating if curiosity can improve sample efficiency in real life.

[5] **SQUIRL: Robust and Efficient Learning from Video Demonstration of Long-Horizon Robotic Manipulation Tasks** NY, USA 11/2019 – 05/2020

Advised by Dr. Peter Allen at Columbia CS

PROBLEM While behavior cloning (BC) is sample-efficient, compounding errors often drive the agent far away from expert demonstrations. In contrast, inverse RL directly recovers the expert reward function at the cost of many real-world trial-and-error experiences. These challenges are amplified when operating with long-horizon image observations.

ALGORITHM Using my Seed-UR5 teleoperation system, we first bootstrapped the learning of a task-conditioned policy with BC. Under this policy, we collect real-world robot trajectories. Finally, we recovered the Q-function from aggregated expert and robot experiences to improve over BC. The refined policy performs a new but related pick-pour-place task from a single demonstration, which we encode as a task embedding.

RESULTS I conducted real-world robot experiments that demonstrate our algorithm outperforms BC with an over 90% success rate on new tasks without trial-and-error at test time.

PUBLICATION Bohan Wu, Feng Xu, Zhanpeng He, **Abhi Gupta**, and Peter Allen. "SQUIRL: Robust and Efficient Learning from Video Demonstration of Long-Horizon Robotic Manipulation Tasks." **IROS 2020.**

[6] **Generative Attention Learning: A "GenerAL" framework for High-Performance Multi-fingered Grasping in Clutter** NY, USA 01/2019 – 08/2019

Advised by Dr. Peter Allen at Columbia CS

PROBLEM Advances in deep learning have motivated data-driven grasping. The supervised setting assumes labels can be generated with top-down grasps. However, these algorithms do not generalize well because the most appropriate grasp for an object surrounded by clutter need not be among the generated labels. Limited to low degree-of-freedom (DOF) manipulators because of high sample-complexity, these strategies are ineffective where greater grasping dexterity and precision is needed.

ALGORITHM We recognized the search space of optimal grasps is limited to the point cloud of the cluttered scene. Our RL formulation for high DOF grasping operates directly in pixel space where the agent learns to attend to local subregions of the scene to find high quality grasps. I observed stable grasps can often be formed with control over a few DOFs.

EXPERIMENTS I deployed our policy optimized with Proximal Policy Optimization in simulation on the UR5-Seed robot. To overcome the mechanical limitations of the hardware, I introduced a new finger-closing algorithm for the anthropomorphic Seed hand. Our real-world setup closely matched the training environment and effortlessly overcame the sim2real gap.

RESULTS The method achieved over 90% grasp success rate across multiple manipulators in unseen single-object and cluttered scenes with depth images from a fixed-viewpoint camera.

PUBLICATION Bohan Wu, Iretiayo Akinola, **Abhi Gupta**, Feng Xu, Jake Varley, David Watkins-Valls, and Peter Allen. "Generative Attention Learning: A General Framework for High-Performance Multi-fingered Grasping in Clutter," **Autonomous Robots, pp. 1–20, 2020**

[7] Electromyography-Driven Hand Teleoperation — NY, USA

Advised by Dr. Peter Allen at Columbia CS — 09/2018 – 05/2019

PROBLEM Robotic teleoperation requires a suite of sensors to track the teleoperator's hand. Brain-computer interfaces provide a single and universal interface for control, but they rely on raw and noisy electromyography (EMG) signals that limit their use to gesture classification. For this reason, the teleoperator can at best execute motion primitives like open or close but rarely anything in-between.

PARTNERSHIP Ctrl-Labs, now part of Facebook Reality Labs, has developed proprietary models capable of predicting continuous-valued hand joint angles from raw EMG with unprecedented accuracy. I introduced a personalization step, where the user collects their own ground truth joint angle data with a dataglove, to further refine these general models.

RESULTS The resulting EMG interface can track the teleoperator's hand to continuously control high DOF manipulators. I showed a simple linear transformation is enough to map 21 DOFs on the human hand to the 5 on the anthropomorphic, under-actuated Seed hand. With this setup, I enabled the user to pick and place objects, such as a mug or can, in the real world.

PRESENTATION [1] Demo. New England Manipulation Symposium, Columbia University. June 2019.
[2] Poster. Data Science Day, Columbia University. April 2019.
[3] Slides. Humanoid Robots Graduate Course, Columbia University. February 2019.

[8] Visual-Tactile Geometric Reasoning — NY, USA

Advised by Dr. Peter Allen at Columbia CS — 05/2018 – 09/2018

PROBLEM Often, the failure cases of even the best grasp planners arise from the object lying partially outside the space enclosed by the manipulator. Tactile sensors can provide feedback to perturb the end-effector pose and adjust grip around the object in hand

EXPERIMENTS At the New England Manipulation Symposium at Yale University, I fabricated the Model-O OpenHand—a three fingered robotic manipulator. By adding a tactile sensor to each finger, I successfully used trilateration, the same way GPS compute position, to localize an object within the manipulator's grasp.

PRESENTATION [1] Demo. New England Manipulation Symposium, Yale University. June 2018.
[2] Poster. Data Science Day, Columbia University. September 2018.
[3] Workshop. Data Science Scholars Bootcamp, Data Science Institute, Columbia University. May 2018.
https://datascience.columbia.edu/news/2018/data-science-scholars-bootcamp/

Work Experience

[1] **Ctrl-Labs (now Facebook Reality Labs)** NY, USA

Science & Interactions Intern, Supervised by Dr. Andrew Berenzweig 05/2019 – 09/2019

Electromyography (EMG) signals are generated by the muscles in our arms when we move our hands. I collected company-wide performance metrics of EMG models on a series of control tasks. By analyzing where these models perform worse on EMG from novel users, I was able to design a calibration procedure to fine-tune two-dimensional cursor control to the specific user.

[2] **Sploreguide** NY, USA

Mobile Development Lead, Founding Member of Tourism Start-up 05/2018 – 09/2018

I developed a universal platform for both tourists and locals to connect on niche experiences around town that are not easily discoverable with tools like Yelp. I rebranded the start-up with a new user interface and user experience that encouraged artists and freelancers to join our community.

[3] **Columbia Daily Spectator** NY, USA

SpecTech Product Team, Web Development 09/2017 – 09/2018

As part of my required work-study on campus, I introduced new products like course scheduling tools to the Columbia community. I co-developed CU@Oracle and webscraped content found on deprecated websites like EZA@CU.

[4] **Brave Control Solutions** Windsor, ON, Canada

Software Engineering Intern 06/2016 – 08/2016

I optimized a routing algorithm for an automated parking garage. I also designed 3D simulation environments where robots and manufacturing processes could be rigorously tested for defects.

[5] **City Center Walk-in Medical Clinic** Windsor, ON, Canada

Software Engineering Intern 06/2015 – 08/2015

Patient medical data in privately-run walk-in clinics is often disorganized and stored on shelves. I developed a custom electronic medical recording (EMR) system to improve the daily workflow of doctors and nurses. I also built a real-time display of the current estimated waiting time per patient.

Selected Projects in Machine Learning

[1] **Reinforcement Learning Playground** 07/2020 – 09/2020

I developed and unit-tested a library of model-free and model-based reinforcement learning algorithms, including REINFORCE, PPO, TRPO, SAC, SARSA, DQN, DDQN, ME-TRPO and MBPO. The simplicity of the framework encourages students to learn RL and researchers to quickly experiment with new ideas. The modularity of the framework allows anybody to integrate RL into their own setup and plug-and-play with both gym and user-defined environments.

[2] **TrackIt! Self-supervised Motion Tracking in 2D Environments** 09/2019 – 12/2019

TrackIt! is a self-supervised framework for tracking the motion of perfectly elastic collisions in a two-dimensional physics environment from pixel observations. I explored a variety of neural network architectures like autoencoders, variational autoencoders, LSTMs, and ConvLSTMs to predict image frames in the future from the past.

The Medium blogpost is available: https://medium.com/@asg2233/trackit-9f9c17183838

[3] **Learning from EMG Synergies to Grasp Objects by Superquadric Representation** 01/2019 – 05/2019

I collected a dataset of human grasps on a variety of YCB objects, including raw EMG signals and hand joint angles. These objects were simple enough to parametrize with the model of a superquadric. With a deep autoencoder, I was able to encode these synergies into a latent space that could be used to directly learn a grasping policy with reinforcement learning. The poster is available here

Hackathons

[1] **DevFest** Columbia University, NY, USA

Abhi Gupta and Jaidev Shah 02/2018 – 04/2018

Project I led Audiv, a multimodal speech transcription and data visualization platform. I launched a bot service within Facebook group chats to transcribe speech to text in real-time. This was challenging because I had to bypass Facebook security rules.

Prizes [1] Top 5 Hacks
[2] Qualified as Semi-Finalists to Columbia's Venture Competition

[2] **Facebook World Hackathon Finals** Menlo Park, CA, USA

Abhi Gupta and Atilla Saadat 10/2016 – 11/2016

Project I developed Watson, a cellular personal assistant with all of the functionality of Siri without the need for a network connection. This required redirecting SMS messages to a NoSQL database and improving Facebook's NLP API wit.ai and backend service Parse Core.

Prizes [1] Top 4 Hacks.
[2] An all-expenses trip for several weeks at the Facebook Headquarters
[3] Because I was a minor at the time, my mom was offered separate travel accommodations while at Facebook so I could participate in the hackathon.

Award [1] "Youngest Finalist at the World Hackathon Finals." I was 15 years old and the only high school student to ever qualify for this collegiate event.

[3] **MHacks 6** University of Michigan, MI, USA

Abhi Gupta and Atilla Saadat 08/2016 – 09/2016

Project I invented Gateway, the first iOS messaging app of its kind, to provide bot functionality. I rebuilt and integrated the iOS Messages user interface into the app from scratch. Out of request, I developed a desktop version of the platform as well.

Prizes [1] Top 5 Hacks
[2] "Best Use of Facebook Parse," an all-expenses trip to compete in the Facebook World Hackathon Finals, and a $500 Amazon Gift Card, sponsored by Facebook.
[3] "Best iOS App at Hackathon," Apple Watch, and an invitation to Cupertino Headquarters, sponsored by Apple.
[4] $500 Amazon Gift card and an invitation to Palo Alto Research Center, sponsored by Ford Research and Development
[5] "Best Use of the Expedia API", $150 travel voucher, and a suitcase, sponsored by Expedia.

News For taking over the hackathon with my success as a 15-year-old, I was invited to an exclusive dinner with Apple engineers and recruiters, to Apple's One Infinity Loop campus in Cupertino, and to Ford R&D to explore internship opportunities.

Science Fairs

[1] **Eye-Tracking for Patients with MS, PPMS, and ALS** Windsor, ON, Canada

Canada-Wide Science Fair 2014 05/2014 – 06/2014

Project People with paralysis struggle to use existing communication aids because of their limited functionality and high-end price tag. I created a cheap $150 alternative that uses off-the-shelf parts like a PlayStation camera to track the user's gaze with 95.3% patient-tested accuracy. I developed the tracking software with OpenCV and prototyped my own hardware and circuit boards. My device allowed both an ALS patient and a quadriplegic to communicate with their families for the first time.

National Awards
[1] Silver Medal – Second Place in the Canada-Wide Science Fair.
[2] Western University Entrance Scholarship of $2000

Regional Awards
[1] Gold Medal - Best Project at Windsor Regional Science & Technology Engineering Fair.
[2] An-Noor Innovation Award and $100 for the Best Engineering Project that will enhance the quality of life for all humanity.
[3] Engineering Award, a plaque and $100 for the Best Engineering Project, sponsored by Professional Engineers Ontario - Windsor-Essex.
[4] $2000 in funding to cover all expenses to represent the city at the National Science Fair, sponsored by the Greater Essex County District School Board
[5] Promising Young Engineer Award, a $100 Bursary and a medal, sponsored by the University of Windsor
[6] $1000 Entrance Scholarship to the University of Windsor for the Best Overall Project
[7] Windsor Islamic Association Award of Excellence and a Gift Certificate
[8] Windsor Islamic Association Award of Merit and a Gift Certificate
[9] Windsor-Essex Chapter of Ontario Association of Certified Engineering Technicians and Technologists Award, medal and cash prize

[2] **Improving the Efficiency of Photovoltaic Solar Panels** Windsor, ON, Canada

Canada-Wide Science Fair 2013 05/2013 – 06/2013

Project The amount of energy produced by photovoltaic panels is limited due to their immobility. When panels maximize their exposure to the sun, their efficiency increases significantly. Motivated by the sunflower, a plant that continually turns with the sun through photosynthesis, I engineered a similar mechanism so solar panels can reorient themselves throughout the day without expelling much energy.

National Awards
[1] Bronze Medal – Third Place in the Canada-Wide Science Fair
[2] Innovation in Renewable Energy Award and a $500 cash prize
[3] Western University Entrance Scholarship of $1000

Regional Awards
[1] Gold Medal - Best Junior Project at Windsor Regional Science Fair
[2] Windsor-Essex Chapter of Ontario Association of Certified Engineering Technicians and Technologists Award.
[3] $2000 in funding to cover all expenses to represent the city at the National Science Fair, sponsored by the St. Clair College

Relevant Coursework

Mathematics

Calculus 2, Multivariable Calculus, Linear Algebra, Probability and Statistics, Optimization Models and Methods, Advanced Optimization, Stochastic Models, Signals and Systems

Computer Science

Data Structures, Artificial Intelligence, Deep Learning for Computer Vision, Computational Aspects of Robotics, Humanoid Robots, Robot Learning, Reinforcement Learning

Software Engineering

Databases, Cloud Computing and Microservices, Advanced Programming

Technical Skills

Web Development

JavaScript, TypeScript, HTML, CSS, Node.js, NoSQL, SQL, AWS, React

Mobile Development

Swift, Objective-C, Xcode

Software Engineering

Python, Java, C, C++, ROS, Git, Unix, Bash

Machine Learning

NumPy, SciPy, PyTorch, TensorFlow, Pyro, OpenCV, PyTorch 3D, PyBullet, Mujoco

References

[1] **Dr. Shipra Agrawal, Agrawal Research Group at Columbia University**
Cyrus Derman Assistant Professor of Industrial Engineering and Operations Research
Relationship: Research Advisor

[2] **Dr. Hod Lipson, Creative Machines Lab at Columbia University**
Professor of Mechanical Engineering
Relationship: Research Advisor

[3] **Dr. Shuran Song, Columbia RoboVision Lab at Columbia University**
Assistant Professor of Computer Science
Relationship: Research Advisor

[4] **Dr. Donald Ferguson, Former CTO of Dell, Columbia University**
Professor of Professional Practice in Computer Science
Relationship: Mentor