

# Anuj Mahajan

Apple, Cupertino, USA  
✉ [anuj.mahajan.phd@gmail.com](mailto:anuj.mahajan.phd@gmail.com)  
📄 [Anuj-Mahajan.github.io](https://Anuj-Mahajan.github.io)



## Research Interests

Artificial Intelligence (AI), Machine Learning (ML), Generative AI, Reinforcement Learning, LLM Agents & Multi-modal Planning/Reasoning, Natural Language Processing, Multi-Agent Systems, AI & Generalization, Computational Learning Theory, Safety & Alignment in AI, Large-scale AI, Optimization

## Education

- 2017–2022 **Doctor of Philosophy in Computer Science**, *University of Oxford*, U.K.,  
**Ph.D. advisor:** Prof. Shimon Whiteson, **Thesis:** Reinforcement Learning in Large State-Action Spaces.
- 2011–2016 **Master of Technology in Computer Science & Engg (Dual degree)**, *Indian Institute of Technology*, Delhi, **Thesis** : Exploring new techniques for MAP Inference in Markov Random Fields.
- 2011–2016 **Bachelor of Technology in Computer Science & Engg (Dual degree)**, *Indian Institute of Technology*, Delhi.

## Work Experience

- 2024-Current **Senior Research Scientist**, *Apple*, Cupertino, USA.  
Working in the Machine Learning Foundations team at the intersection of the following areas:
- Foundation Models / Generative AI
  - Reinforcement Learning
  - LLM Agents and Multimodal Planning/Reasoning
- My work explores the spectrum of decision-making methods as well as how to effectively combine them to build efficient, robust agents that generalize well across new tasks.
- 2023-2024 **Applied Scientist II, Reinforcement Learning team lead**, *Amazon*, SF, USA.  
Working in the Amazon Ads team at the intersection of the following areas: Generative AI, Reinforcement Learning (RLHF/RLAIF, LLM Agents), Natural Language Processing, Computer Vision.
- 2021 **Research Scientist Intern**, *DeepMind*, London, UK.  
Open Ended Learning Systems: Creating foundational AI agents for large task spaces.
- 2020-2021 **Research Scientist Intern**, *J.P. Morgan Chase*, London, UK.  
Safe Reinforcement Learning for long term decision making with constraints.
- 2019-2020 **Research Scientist Intern**, *NVIDIA*, Santa Clara, USA.  
Multi-Agent Reinforcement Learning using tensorised function approximations.
- 2016-2017 **Research Scientist**, *Xerox Research Centre*.  
Worked in the Machine Learning and Statistics Group in the following areas:
- Deep Reinforcement Learning
  - Probabilistic Graphical Models
  - Ranking for Duelling Bandits
- 2014 **Research Intern**, *Xerox Research Centre*.  
Feature selection methods using Wavelet Packet transforms, published in CoNLL 2015.

---

## Publications<sup>+</sup>

### Conference/Journals/Preprints

Siddharth Nayak, Adelmo Morrison Orozco, Marina Ten Have, Vittal Thirumalai, Jackson Zhang, Darren Chen, Aditya Kapoor, Eric Robinson, Karthik Gopalakrishnan, Brian Ichter, James Harrison, Anuj Mahajan, and Hamsa Balakrishnan. Long-horizon planning for multi-agent robots in partially observable environments. In *Advances in Neural Information Processing Systems*. 2024 [**NeurIPS**].

Anuj Mahajan and Amy Zhang. Generalization across observation shifts in reinforcement learning. 2023 [**arXiv**].

Anuj Mahajan, Mikayel Samvelyan, Lei Mao, Viktor Makoviychuk, Animesh Garg, Jean Kossaifi, Shimon Whiteson, Yuke Zhu, and A Anandkumar. TESSERACT: Tensorised actors for multi-agent reinforcement learning. In *Thirty-eighth International Conference on Machine Learning*. 2021 [**ICML**].

Tarun Gupta, Anuj Mahajan, Bei Peng, Wendelin Boehmer, and Shimon Whiteson. UNEVEN: Universal value exploration for multi-agent reinforcement learning. In *Thirty-eighth International Conference on Machine Learning*. 2021 [**ICML**].

Adam Stooke, Anuj Mahajan, Catarina Barros, Charlie Deck, Jakob Bauer, Jakub Sygnowski, Maja Trebacz, Max Jaderberg, Michael Mathieu, Nat McAleese, Nathalie Bradley-Schmieg, Nathaniel Wong, Nicolas Porcel, Roberta Raileanu, Steph Hughes-Fitt, Valentin Dalibard, and Wojciech Marian Czarnecki. Open-ended learning leads to generally capable agents. 2021 [**DeepMind Tech report**].

Anuj Mahajan\*, Matthew Fellows\*, Tim GJ Rudner, and Shimon Whiteson. VIREL: A variational inference framework for reinforcement learning. In *Thirty-third Conference on Neural Information Processing Systems*. 2019 [**Spotlight, NeurIPS**].

Anuj Mahajan, Tabish Rashid, Mikayel Samvelyan, and Shimon Whiteson. MAVEN: Multi-agent variational exploration. In *Thirty-third Conference on Neural Information Processing Systems*. 2019 [**NeurIPS**].

Anuj Mahajan, Mikayel Samvelyan, Tarun Gupta, Benjamin Ellis, Mingfei Sun, Tim Rocktäschel, and Shimon Whiteson. Generalization in cooperative multi-agent systems. 2022 [**arXiv**].

Tonghan Wang, Tarun Gupta, Anuj Mahajan, Bei Peng, Shimon Whiteson, and Chongjie Zhang. Rode: Learning roles to decompose multi-agent tasks. In *Ninth International Conference on Learning Representations*. 2021 [**ICLR**].

Anuj Mahajan and Theja Tulabandhula. Symmetry detection and exploitation for function approximation in deep RL. In *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems*. International Foundation for Autonomous Agents and Multiagent Systems, 2017 [**AAMAS**].

Happy Mittal, Anuj Mahajan, Vibhav G Gogate, and Parag Singla. Lifted inference rules with constraints. In *Advances in Neural Information Processing Systems 28*, pages 3501–3509. Curran Associates, Inc., 2015 [**NeurIPS**].

Mingfei Sun, Ben Ellis, Anuj Mahajan, Sam Devlin, Katja Hofmann, and Shimon Whiteson. Trust-region-free policy optimization for stochastic policies. 2022 [**AAMAS**].

Anuj Mahajan, Sharmistha Jat, and Shourya Roy. Feature selection for short text classification using wavelet packet transform. In *Proceedings of the Nineteenth Conference on Computational Natural Language Learning*, pages 321–326. Association for Computational Linguistics, 2015 [**CoNLL**].

---

<sup>+</sup> Full updated list available at Google Scholar: [Here](#)

Jerry Zhang, Eldon Schoop, Jeff Nichols, Anuj Mahajan, and Amanda Swearngin. From interaction to impact: Towards safer ai agents through understanding and evaluating ui operation impacts. 2024 [**arXiv**].

Benjamin Ellis, Skander Moalla, Mikayel Samvelyan, Mingfei Sun, Anuj Mahajan, Jakob Foerster, and Shimon Whiteson. Smacv2: A new benchmark for cooperative multi-agent reinforcement learning. In *Conference on Neural Information Processing Systems*. 2023 [**NeurIPS**].

Kinal Mehta, Anuj Mahajan, and Pawan Kumar. Effects of spectral normalization in multi-agent reinforcement learning. In *IEEE International Joint Conference on Neural Networks*. 2023 [**IJCNN**].

Mingfei Sun, Anuj Mahajan, Katja Hofmann, and Shimon Whiteson. Softdice for imitation learning: Rethinking off-policy distribution matching. 2021 [**arXiv**].

Anuj Mahajan and Theja Tulabandhula. Symmetry learning for function approximation in reinforcement learning. 2017 [**arXiv**].

#### Workshops

Anuj Mahajan and Amy Zhang. Conditional bisimulation for generalization in reinforcement learning. In *Topology, Algebra, and Geometry in Machine Learning (TAG-ML)*. 2023 [**ICML**].

Anuj Mahajan, Mikayel Samvelyan, Lei Mao, Viktor Makoviychuk, Animesh Garg, Jean Kossaifi, Shimon Whiteson, Yuke Zhu, and A Anandkumar. Reinforcement learning in factored action spaces using tensor decompositions. In *Quantum Tensor Networks in Machine Learning Workshop*. 2021 [**NeurIPS**].

Pascal Van Der Vaart, Anuj Mahajan, and Shimon Whiteson. Model based multi-agent reinforcement learning with tensor decompositions. In *Quantum Tensor Networks in Machine Learning Workshop*. 2021 [**NeurIPS**].

Luisa Zintgraf, Maximilian Igl, Kyriacos Shiarlis, Anuj Mahajan, Katja Hofmann, and Shimon Whiteson. Variational task embeddings for fast adaptation in deep reinforcement learning. In *Structure & Priors in RL Workshop*. 2019 [**ICLR**].

Anuj Mahajan and Theja Tulabandhula. Discovering symmetries for sample efficient reinforcement learning. In *The Multi-disciplinary Conference on Reinforcement Learning and Decision Making*. 2017 [**RLDM**].

Siddharth Nayak, Adelmo Morrison Orozco, Marina Ten Have, Vittal Thirumalai, Jackson Zhang, Darren Chen, Aditya Kapoor, Eric Robinson, Karthik Gopalakrishnan, Brian Ichter, James Harrison, Anuj Mahajan, and Hamsa Balakrishnan. Map-thor: Benchmarking long-horizon multi-agent planning frameworks in partially observable environments. In *Multi-modal Foundation Model meets Embodied AI*. 2024 [**ICML**].

---

#### Patents

USA Method and system for predicting requirements of a user for resources over a computer network, Number: US010417578B2

USA Personalizing application interfaces based on usage, Number: US011112950B2

---

#### Teaching

2019 **Tutor**.

Tutor for Machine learning for Computer Science & Philosophy undergrads, Trinity term, Hertford College, University of Oxford.

2019 **Teaching Assistant**.

TA for Reinforcement Learning, Hilary term, Autonomous Intelligent Machines and Systems (AIMS), University of Oxford.

---

\* Equal contribution

## 2015-2016 **Teaching Assistant.**

TA for the following courses at IIT, Delhi:

- Machine Learning (COL774) Spring semester 2015-16.
- Computer Networks (COL334) Fall semester 2015-16.

---

## Reviewing & Program Committee

**NeurIPS** Neural Information Processing Systems, 2019, 2020, 2021, 2022, 2023

**ICML** International Conference on Machine Learning, 2021, 2023

**AISTATS** Artificial Intelligence and Statistics, 2021

**ICLR** International Conference on Learning Representations, 2021, 2023

**JMLR** Journal of Machine Learning Research, 2020

**AAAI** Association for the Advancement of Artificial Intelligence, 2023

**IEEE** IEEE Transactions on Neural Networks and Learning Systems, 2022

**ELEC** Electronic Commerce Research, Springer, 2018, 2022

**TMLR** Transactions on Machine Learning Research, 2022

**NeurIPS** Quantum Tensor Networks in Machine Learning Workshop, 2021

**NeurIPS** Deep Reinforcement Learning Workshop, 2022

**IEEE** IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022

---

## Technical skills

Python, Java, C/C++, Prolog, SQL, Ocaml, Assembly

Pytorch, Tensor Flow, Jax, Docker, Matlab, Android, Eigen, AWS, Huggingface

---

## Relevant Courses

Advanced Machine Learning, Computational Learning theory, Machine Learning, Probabilistic Graphical Models, Adv. Algorithms, Data Mining, Computer Vision, Theory of Computation, Computational Biology, Molecular Cell Biology

---

## Scholarships

- Awarded J.P. Morgan AI fellowship 2020.
- Awarded IBM PhD fellowship 2020 (declined).
- Awarded Google Deepmind Scholarship 2017-20 for doctoral studies at University of Oxford.
- Awarded Drapers Hertford graduate Scholarship 2017-20 for doctoral studies at University of Oxford.
- Awarded Microsoft Student Travel Grant for presenting research paper at CoNLL 2015, Beijing, China.
- Awarded Microsoft Student Travel Grant for presenting research paper at NeurIPS 2015, Montreal, Canada.
- Kishore Vaigyanic Protsahan Yojana(KVPY) fellowship awarded by the Department of Science and Technology, Government of India. (**Given to 200 fellows chosen from around one million applicants**)
- Awarded Indian National Association of Engineers (INAE) grant 2015.
- National Talent Search Examination(NTSE) fellowship awarded by NCERT, Department of Education, Government of India. (**500 scholars chosen from around one million applicants**)

---

## Awards & Achievements

- Uber AI resident 2020 (Program rescinded due to covid19)
- Indian Institute of Technology, Delhi, Institute Merit Award : Received the prestigious IITD merit award given to **top 5% students in the institute.**

- Winner, Microsoft 'code.fun.do' : Programming event organized by Microsoft on 16-17/02/2013
- Won the Award of Excellence in Australian National Chemistry Quiz(ANCQ) for securing **All India Rank - 1** for three consecutive years (2006-08)
- Represented the state at Indian National Mathematics Olympiad and Astronomy Olympiad.
- Secured 8th position in the Regional Mathematical Olympiad, 2008 organized by NBHM, Government of India.
- Best Research Poster award at the Xerox open house 2014 poster presentation event.

---

## Media

- MIT Technology Review on Open Ended Learning: [Link](#)
- DeepMind blog on Generally Capable Agents: [Link](#)
- AI plays catch, Two Minute Papers: [Link](#)
- Multi-Agent Perspective to AI, talk at GoodAI: [Link](#)