Statement of Purpose

Aditya Makkar

This statement is intended to convey my desire to pursue a Ph.D. in the area of multi-agent reinforcement learning (MARL) at New York University under the guidance of Professor Eugene Vinitsky. MARL is a highly interdisciplinary subject that builds upon ideas from game theory, control theory, learning theory, optimization, statistics and probability theory in a unique way to approach the important and challenging problem of designing autonomous agents. I would like to build a research career in furthering our theoretical understanding of this exciting area and applying this understanding to solve practical problems in designing autonomous agents making robust sequential decisions in complex environments, like in autonomous driving.

Background and Research Experience

I developed a love for mathematics and an appreciation for elegant arguments in high school by participating in math olympiads. I got to pursue topics like number theory, graph theory, combinatorics, functional equations, inequalities and geometry beyond the prescribed syllabus. I was an **Indian National Mathematical Olympiad finalist** the next two years (2010 and 2011), and came 3$^{\text{rd}}$ in the state of Uttar Pradesh during my final year. I cleared the Joint Entrance Examination and gained admission into the prestigious Indian Institute of Technology Delhi, where I majored in mechanical engineering. I received the **semester merit award** twice for academic performance (a GPA of 9.7/10 and department topper). I took many computer science and mathematics courses, and an interest in machine learning took root.

During my undergrad I interned at **Goldman Sachs** in a team researching and developing **machine learning** and statistical tools for surveillance models which must process terabyte-scale data and flag anomalous activities. I received a full-time offer from the team and I worked here for three years with my manager Dr. Howard Karloff on a variety of projects, e.g., text summarization using clustering algorithms (like affinity propagation, hierarchical clustering and k-means), word2vec , doc2vec, fuzzy name matching using edit distance, anomaly detection using variational autoencoders, active learning, and Latent Dirichlet Allocation (LDA), the last two done in collaboration with Professor John Paisley from Columbia University. These projects helped me develop strong programming skills and an understanding of the complexities of dealing with real, as opposed to toy, data. With the help of my manager, I open-sourced a MapReduce implementation of word2vec. It can be found on GitHub [MK19].

Motivated to pursue research and develop a deeper understanding of machine learning I joined the Ph.D. program at Columbia advised by Professor Paisley. I worked on developing a **Bayesian nonparametric ensemble** method with my advisor and a team in the environmental sciences. The project was motivated by the practical problem of predicting $\mathrm{PM}_{2.5}$ (particulate matter found in the air that are 2.5 μm or smaller) concentration in the United States by combining several existing models. An ensemble is typically of the form $\widehat{S} := f(S_1, \ldots, S_L) \colon \mathscr{X} \to \mathscr{Y}$, $f(S_1, \ldots, S_L) = \sum_{\ell=1}^{L} c_\ell S_\ell$, where $\mathscr{X} \subset \mathbb{R}^3$ (latitude, longitude, time), $\mathscr{Y} = [0, \infty)$, $\{S_1, \ldots, S_L\}$ are models to be combined, each one a mapping from $\mathscr{X}$ to $\mathscr{Y}$, and $\{c_1, \ldots, c_L\}$ are fixed weights that sum to 1. Our main goals in this project were having an adaptive notion of weights, i.e., each $c_\ell \colon \mathscr{X} \to [0,1]$ is a function of the input space, and providing uncertainty estimates for the ensemble prediction. I contributed by 1. significantly generalizing this basic framework by using dependent tail-free processes [JH11] as priors on the weights c_ℓ's to model hierarchical dependence, 2. using random Fourier features [RR08], which allowed this model to scale up from city level to global level, and 3. implementing it from scratch and handling the related programming challenging. Please see our paper [MLR$^+$22] for more details. Dealing with real data and trying to model it taught me how to run an iterative loop of better generalizations.

The expressive capabilities of kernel methods and nonparametric statistics led me to explore how to combine them. I worked on developing a **minimax optimal test statistic for the problem of goodness of fit**, where we are given a sample of i.i.d. observations, and the goal is to test if they come from some fixed known probability distribution. Given a measurable space $(\mathscr{X}, \mathfrak{X})$, a symmetric positive definite kernel $k \colon \mathscr{X} \times \mathscr{X} \to \mathbb{R}$ and its associated reproducing kernel Hilbert space $\mathscr{H}_k$, we can map probability measures on $(\mathscr{X}, \mathfrak{X})$ to $\mathscr{H}_k$ using $\mathbb{P} \mapsto \mu_{\mathbb{P}}(\cdot) := \int_{\mathscr{X}} k(\cdot, x)\, \mathrm{d}\mathbb{P}(x)$. This gives us a notion of distance between probability measures by using the distance between, so called, *mean elements* $\mu_{\mathbb{P}}$. Using results from operator analysis I tried to construct a U-statistic statistic whose rate of convergence is minimax optimal. Although I failed to find such a statistic, the project taught me how to do theoretical research and I learnt the importance of good advicing.

I developed a strong foundation by taking many **mathematics courses** and participating in multiple reading groups and seminars at Columbia. For example, I have taken courses in abstract analysis, probability theory, topology, abstract algebra, high-dimensional statistics, the concentration of measure and empirical process theory, and participated in reading groups for stochastic analysis, ergodic theory, optimal stopping theory and optimal control. I also enjoy teaching and have been a **teaching assistant** for a Masters level course on introduction to machine learning twice and a Masters level course digital signal processing once.

After leaving the Ph.D. program, I have been working at **Morgan Stanley** as a quantitative strategist at the interest rates trading desk, where I design tools for traders which they can use to price the securities, hedge the risks and get recommendations for trades.

Why and How?

A Ph.D. has a substantial opportunity cost, but for me it is a thoughtful conscious choice. At the most primitive level my **reasons to pursue a Ph.D.** stem from a desire to pursue knowledge for its own sake and get a deep understanding of machine learning by creating new theory and engineering solutions. Theory is not about simply proving theorems, but rather about providing a framework to make difficult things easy. It is with this perspective that I want to work on integrating theoretical research with application-driven research. This idée fixe to become the best intellectual version of myself is grounded in practical realization that I enjoy research and I want to build a career in research. And therefore, as Weil said, I would like to abandon myself to the selfish joys of creative work [Wei50].

I immediately fell in love with the mathematics of optimal stopping and optimal control of stochastic processes (in the spirit of [Nev75] (Chapter IV) and [CRS71, CD96, Shi19, PS06, FR75, FS05]) that I got introduced to in classes and reading groups organized by Professor Ioannis Karatzas. The topic of **Markov decision processes** has fascinated me ever since, and I gave two talks on the foundations of this topic from [BS78, HLL96] (see my blog post [Mak22] giving a general framework). Not only are there many challenging analysis and probability theory questions (bringing to the surface deep results like selection theorems [Wag77]), these topics form the foundation for reinforcement learning.

There are many **research directions** I would like to work on given my background and Professor Vinitsky's expertise. An important need is dealing with high-dimensional data and designing MARL algorithms for more general continuous spaces. Interesting and challenging problems (like convergence of algorithms in games) lie at the connection between discretization of continuous dynamics and the continuous dynamics themselves. The need to study this is immediately apparent once one realizes that a continuous time Markov process stops giving us Markov observations once we discretize the time. There's also the need to study objective functions other than discounted rewards; for example, the long-run average expected reward [ABFG+93].

The challenge of **designing practical algorithms** is that reality may not respect artificial assumptions of smoothness and niceness imposed in theory. And therefore, developing robust engineering solutions is another important topic I would like to work on. The subject of MARL provides a convenient framework to develop complex behavioural agents, like in autonomous vehicles and robotics. I would love to work on creating algorithms to tackle dynamic, non-stationary environments. Creative ideas in recent papers like [HLC+21] (a MARL algorithm that allows agents to act optimally without assuming any shared conventions — zero-shot coordination) and [JWF+22, BWL+23] (a MARL algorithm that allows agents to learn human-like policies while avoiding the failure modes of self-play and imitation learning), sent to me by Professor Vinitsky, motivate me and prove that this is a very active research area.

Given my background and experience, joining the Ph.D. program with Professor Vinitsky would offer me an excellent opportunity to start contributing to this intellectual landscape. His extensive work in this area, especially in developing benchmarks (like in [YVV+22]), provides me with plenty of directions for future research. I particularly like his works on developing robust RL frameworks [VDP+20, DJV+20, MJPH+23], a multi-agent system that allows agents to learn social norms through public sanctions [VKA+23], and using the framework of decentralized MARL to improve the throughput of a traffic bottleneck in a mixed autonomy setting [VLPB23].

References

[ABFG+93] Aristotle Arapostathis, Vivek S. Borkar, Emmanuel Fernández-Gaucherand, Mrinal K. Ghosh, and Steven I. Marcus. Dicrete-time controlled markov processes with average cost criterion: A survey. *SIAM J. Control and Optimization*, 31(2):282–344, 1993.

[BS78] Dimitri P. Bertsekas and Steven E. Shreve. *Stochastic Optimal Control: The Discrete Time Case*, volume 139 of *Mathematics in Science and Engineering*. Academic Press, 1978.

[BWL+23] Anton Bakhtin, David J Wu, Adam Lerer, Jonathan Gray, Athul Paul Jacob, Gabriele Farina, Alexander H Miller, and Noam Brown. Mastering the game of no-press diplomacy via human-regularized reinforcement learning and planning. *International Conference on Learning Representations*, 2023.

[CD96] R. Cairoli and Robert C. Dalang. *Sequential Stochastic Optimization*. Wiley Series in Probability and Statistics. Wiley Publishers, 1996.

[CRS71] Y. S. Chow, Herbert Robbins, and David Siegmund. *Great Expectations: The Theory of Optimal Stopping*. Houghton Mifflin Company, 1971.

[DJV+20] Michael Dennis, Natasha Jaques, Eugene Vinitsky, Alexandre Bayen, Stuart Russell, Andrew Critch, and Sergey Levine. Emergent complexity and zero-shot transfer via unsupervised environment design. *NeurIPS*, 2020.

[FR75] Wendell Fleming and Raymond Rishel. *Deterministic and Stochastic Optimal Control*, volume 1 of *Stochastic Modelling and Applied Probability*. Springer New York, 1975.

[FS05] Wendell Fleming and H.M. Soner. *Controlled Markov Processes and Viscosity Solutions*, volume 25 of *Stochastic Modelling and Applied Probability*. Springer New York, 2005.

[HLC+21] Hengyuan Hu, Adam Lerer, Brandon Cui, Luis Pineda, Noam Brown, and Jakob Foerster. Off-belief learning. *Proceeddings of the 38th International Conference on Machine Learning*, 2021.

[HLL96] Onésimo Hernández-Lerma and Jean Bernard Lasserre. *Discrete-Time Markov Control Processes: Basic Optimality Criteria*, volume 30 of *Stochastic Modelling and Applied Probability*. Springer, 1996.

[JH11] A. Jara and T. E. Hanson. A class of mixtures of dependent tail-free processes. *Biometrika*, 98(3):553–566, 2011.

[JWF+22] Athul Paul Jacob, David J. Wu, Gabriele Farina, Adam Lerer, Hengyuan Hu, Anton Bakhtin, Jacob Andreas, and Noam Brown. Modeling strong and human-like gameplay with kl-regularized search. *Proceeddings of the 38th International Conference on Machine Learning*, 2022.

[Mak22] Aditya Makkar. *Blog post on Markov decision processes*. 2022.

[MJPH+23] Ishita Mediratta, Minqi Jiang, Jack Parker-Holder, Michael Dennis, Eugene Vinitsky, and Tim Rocktäschel. Stabilizing unsupervised environment design with a learned adversary. *Conference on Lifelong Learning Agents (CoLLAs)*, 2023.

[MK19] Aditya Makkar and Howard Karloff. *https://github.com/goldmansachs/MRWord2Vec*. 2019.

[MLR+22] Aditya Makkar, Jeremiah Liu, Sebastian Rowland, Lawrence Chillrud, Marianthi-Anna Kioumourtzoglou, Brent Coull, and John Paisley. *https://makkar.github.io/bne.pdf*. 2022.

[Nev75] J. Neveu. *Discrete-Parameter Martingales*, volume 10 of *North-Holland Mathematical Library*. North-Holland Publishing Company, 1975.

[PS06] Goran Peskir and Albert N. Shiryaev. *Optimal Stopping and Free-Boundary Problems*. Lectures in Mathematics, ETH Zürich. Birkhäuser Basel, 2006.

[RR08] Ali Rahimi and Benjamin Recht. Random features for large-scale kernel machines. In J. Platt, D. Koller, Y. Singer, and S. Roweis, editors, *Advances in Neural Information Processing Systems*, volume 20. Curran Associates, Inc., 2008.

[Shi19] Albert N. Shiryaev. *Stochastic Disorder Problems*, volume 93 of *Probability Theory and Statistical Modelling*. Springer Nature Switzerland, 2019.

[VDP+20] Eugene Vinitsky, Yuqing Du, Kanaad Parvate, Kathy Jang, Pieter Abbeel, and Alexandre Bayen. Robust reinforcement learning using adversarial populations. 2020.

[VKA+23] Eugene Vinitsky, Raphael Köster, John P Agapiou, Edgar A Duéñez-Guzmán, Alexander S Vezhnevets, and Joel Z Leibo. A learning agent that acquires social norms from public sanctions in decentralized multi-agent settings. *Collective Intelligence*, 2023.

[VLPB23] Eugene Vinitsky, Nathan Lichtlé, Kanaad Parvate, and Alexandre Bayen. Optimizing mixed autonomy traffic flow with decentralized autonomous vehicles and multi-agent reinforcement learning. *ACM Transactions on Cyber-Physical Systems*, 2023.

[Wag77] D. H. Wagner. Survey of measurable selection theorems. *SIAM Journal of Control and Optimization*, 15(5):859–903, 1977.

[Wei50] Andre Weil. The future of mathematics. *The American Mathematical Monthly*, 57(5):295–303, 1950.

[YVV+22] Chao Yu, Akash Velu, Eugene Vinitsky, Jiaxuan Gao, Yu Wang, Alexandre Bayen, and Yi Wu. The surprising effectiveness of ppo in cooperative multi-agent games. *NeurIPS Track on Datasets and Benchmarks*, 2022.