Human-level, domain-general policy learning is out of reach of current neural architectures. No individual neural network can match the human ability to reason, satisfy safety constraints, and learn composable skills. This domain-agnostic versatility also demands an invariance to extraneous details, which is difficult to obtain solely from existing methods of statistical generalization. Additional inductive biases appear to be necessary. The goal of my doctoral studies would be to introduce a new kind of inductive bias for **computing with concepts**, improving the sample efficiency of policy learning.

As a Master's student in Professor Shuran Song's Columbia Artificial Intelligence and Robotics Lab[1], I became interested in improved inductive biases for policy learning, particularly for safety constraints. In my Master's thesis, I introduced Koopman Constrained Policy Optimization (KCPO), which I presented at this year's International Conference on Machine Learning [1]. My main contribution is an inductive bias combining the fields of differentiable optimal control and Koopman operator theory. In addition to Professor Song, I was mentored by Brandon Amos[2] and Professor Steve Brunton[3]; the two of them are pioneers in each field. The superior performance of KCPO illustrates that a well-chosen inductive bias can improve sample efficiency and discover behaviors that typical architectures struggle to attain. For KCPO, that behavior is respecting safety constraints, which are crucial in domains like robot surgery and autonomous vehicles, where human well-being is paramount.

My doctoral work would extend my Master's interest in sample-efficient policy learning with my undergraduate training in cognitive science, computer vision, and concept learning. I first explored representation learning at Philips Research, where I implemented neural networks for learning grounded representations that integrate natural language and visual data in a question-answering task. I later explored concept learning in more depth in Professor Nikolaus Kriegeskorte's Visual Inference Lab at Columbia. There, I developed a novel benchmarking platform for interrogating inductive biases of visual concept learning in humans and artificial neural networks [2]. These experiences highlighted the importance of concept learning for sample-efficient, domain-general cognitive systems.

Thus, I aspire to create a **concept computer**: a neural network with an inductive bias for a formal model of computation and disentangled concept learning.[4] This framework would combine symbolic, compositional, and conceptual reasoning with the statistical pattern-matching of neural networks. One promising approach could utilize variational inference, which appears extensively in the disentanglement literature [3] [4], and a Turing-complete Transformer [5] [6]. A Turing-complete policy could synthesize and execute programs by emulating a universal Turing machine. The result would be high-level symbolic reasoning in a discrete action space and low-level control in a continuous action space, representing both percepts and actions as concept vectors.

---

[1] Professor Shuran Song is currently Assistant Professor of Electrical Engineering and, by courtesy, of Computer Science at Stanford University.
[2] Brandon Amos, Ph.D., is a Research Scientist at Meta AI (Fundamental Artificial Intelligence Research).
[3] Professor Steve Brunton is Professor of Mechanical Engineering at the University of Washington, Seattle.
[4] Disentangled representation learning maps each underlying conceptual factor of variation to a specific latent variable factorized from the joint distribution of observed variables. For instance, an agent could play table tennis in zero-shot (domain adaptation), composing previously acquired concepts TABLE and TENNIS.

I would greatly appreciate collaborating with Professor Carl Vondrick, whose Computer Vision II course influenced my understanding of concepts and neural program synthesis. Self-supervised learning, such as Professor Vondrick's work in masked token prediction [7], could be one way to train the concept computer. I also draw inspiration from his ViperGPT, which performs zero-shot visual reasoning via an external API [8]. My proposed concept computer could iterate on this approach with an internal neural API, both synthesizing visual reasoning programs and computing their outputs.

Ultimately, I wish to use reinforcement learning to build a stochastic world model and policy spanning continuous state and action spaces. The concept computer benefits from a virtuous cycle; the agent would use its disentangled concepts to reason and plan its actions, and it would use active perception to facilitate the disentanglement of concepts.

Toward this end, it would be rewarding to work under the mentorship of Professor Shipra Agrawal to build upon the theoretical foundations of reinforcement learning and prove new guarantees surrounding regret minimization, multi-armed bandits, and exploration-exploitation trade-offs for online agents. Her work on bounding regret minimization in multi-armed bandits with Thompson sampling is highly relevant to the concept computer because each reward-maximizing state node from the Turing machine can be formalized as a multi-armed bandit [9].

Stochastic policies are often used in robot learning, but I also have a keen interest in another important domain: systems biology and control of biological systems. I am greatly optimistic about policy learning's potential in a project for live single-cell control. The aim would be rational control of living cells' gene expression. Potential applications include a better understanding of gene regulatory programs as well as directed differentiation. Professor Mohammed AlQuraishi has expressed to me his interest in this project. His research objective and mine align closely because we both aspire to build computational models of biology at an appropriate level of abstraction to illuminate life's causal mechanisms [10]. Furthermore, we both wish to understand the precise causes of the dysregulation associated with disease. I firmly believe that closing the loop of perception and control will further this goal of data-driven modeling of biology. I currently attend Professor AlQuraishi's weekly systems biology lab meetings with great enthusiasm.

I am also drawn to Professor Elham Azizi's research, which has included the application of variational inference to single-cell genomic and epigenomic data. This work involves generative model inference for cell state dynamics in one study [11] and the impact of experimental and sample conditions on cell states in a different study [12]. Live cell control would likely require new methods in generative modeling and single-cell -omics, and I would be grateful to have her guidance.

Through my research experience, I have honed my ability to communicate with scientists of diverse backgrounds and to rigorously examine my ideas for mathematical soundness and empirical performance. My efforts to make complex scientific ideas accessible have not only enhanced my research, but they have also fostered an interest in becoming an effective educator.

During my doctoral studies, I would seek teaching assistant positions to share my passion for discovery with other students.

My goal after the completion of my doctorate would be to become a professional scientist. Columbia's Computer Science Ph.D. Program would be the perfect way to refine my skills and advance the frontier of policy learning with potential applications to computer vision and systems biology.

Bibliography

[1] **Retchin, M.**, Amos, B., Brunton, S., & Song, S. (2023). Koopman Constrained Policy Optimization: A Koopman operator theoretic method for differentiable optimal control in robotics. *International Conference on Machine Learning*. https://differentiable.xyz/papers/paper_45.pdf.

[2] Peters, B., **Retchin, M.**, & Kriegeskorte, N. (2022). Flying objects: Challenging humans and machines in Dynamic Object Vision. *Conference on Cognitive Computational Neuroscience*. https://2022.ccneuro.org/view_paper.php?PaperNum=1301.

[3] Chen, R. T. Q., et al. (2018). Isolating Sources of Disentanglement in Variational Autoencoders. *Neural Information Processing Systems*. https://proceedings.neurips.cc/paper/2018/hash/1ee3dfcd8a0645a25a35977997223d22-Abstract.html.

[4] Higgins, I., et al. (2018). Towards a Definition of Disentangled Representations. https://arxiv.org/abs/1812.02230.

[5] Graves, A., et al. (2016). Hybrid computing using a neural network with dynamic external memory. Nature News. https://www.nature.com/articles/nature20101.

[6] Dehghani, M., et al. (2019) Universal Transformers. *International Conference on Learning Representations*. https://research.google/pubs/pub48062/.

[7] Sun, C., Myers, E., Vondrick, C., et al. (2019). VideoBERT: A Joint Model for Video and Language Representation Learning. *International Conference on Computer Vision*. https://ieeexplore.ieee.org/document/9009570.

[8] Surís, D., Menon, S., & Vondrick, C. (2023). ViperGPT: Visual Inference via Python Execution for Reasoning. https://arxiv.org/abs/2303.08128.

[9] Agrawal, S. & Goyal, N. (2012). Analysis of Thompson Sampling for the Multi-armed Bandit Problem. *JMLR: Workshop and Conference Proceedings*, *23*. http://proceedings.mlr.press/v23/agrawal12/agrawal12.pdf.

[10] AlQuraishi, M. & Sorger, P. (2021). Differentiable biology: using deep learning for biophysics-based and data-driven modeling of molecular mechanisms. *Nature Methods*. https://www.nature.com/articles/s41592-021-01283-4.

[11] Bachireddy, P., Azizi, E., et al. (2021). Mapping the evolution of T cell states during response and resistance to adoptive cellular therapy. *Cell Reports*, *37*(6), 109992. https://www.cell.com/cell-reports/fulltext/S2211-1247(21)01471-6.

[12] Boyeau, P., Azizi, E., et al. (2022). Deep Generative Modeling for Quantifying Sample-Level Heterogeneity in Single-Cell Omics. https://www.biorxiv.org/content/10.1101/2022.10.04.510898v1.