I would like to pursue a PhD in computer science with a focus on artificial intelligence. My long run objective is to understand and create intelligence. In the short run, I am focused on expanding the human-like capabilities of reinforcement learning agents operating in continuous environments, including their ability to develop and translate implicit world knowledge (e.g., in the form of a general value function [5]) into explicit language and back, their ability to provide explicit justification for actions, and their ability to generate and exploit chains of discrete logical reasoning for planning and learning. An intermediate goal is to create an agent that can be taught through verbal instruction in addition to direct experience.

My first substantial research project was on expected utility theory. It started with an undergraduate class on managerial decision making, where we were taught to use expected value over monetary outcomes to evaluate risky decisions. Being a professional poker player, I took issue with the model: it could not capture the idea that one should never bet the bank on a single game. This piqued my curiosity, and I spent some time formalizing my thoughts on the topic. I had to approach several professors with my ideas before professor Archishman Chakraborty, a game theorist, immediately understood the problem I was speaking to. He gave me the statement of the von Neumann-Morgenstern expected utility hypothesis and asked me to prove it. This was followed by several suggested readings and a semester-long independent study on decision making under uncertainty, which I pursued vigorously. One thing I greatly appreciated about this experience was the autonomy: I was able to discover the subject matter in my own way, to a level far beyond what one encounters in the typical classroom.

After my undergraduate studies, I went to law school, where I worked on two independent research projects in the law and economics sphere.[1] One was on optimal design of corporate policy on takeover defenses. The other was on optimal design of a remedial mechanism for governing international trade relations. The common thread between these projects and my project on expected utility theory is *efficient design of agency* (be it that of a human, a firm or a government). During law school, I expressed some interest in pursuing an economics PhD along these lines. Practical considerations held me back; in particular, I had doubts about the impact that normative analysis could have on actual behavior. Instead, I moved to New York to practice corporate law.

While a lawyer, I noticed several gaps in legal technology and, intent on filling them, I began to study computer programming and build software in my off time. This software, some of which is still in use today, has saved hundreds of billable hours. My most recent work, which resulted in a conference publication and best student paper award at ICAIL 2017 [2], is a prototype of an efficient, similarity-based passage retrieval system for legal contracts. It was developed after digesting two information retrieval textbooks and more than 40 research papers, and applies existing search techniques, augmented with a few of my own novelties. As part of this project, I rebuilt the backend (i.e., the format of the inverted index, and the way documents are added, stored and retrieved) of the popular Python search framework "Whoosh", more than doubling its indexing and retrieval speed and allowing it to handle the relatively big data I performed my experiments on (12GB raw / 2GB indexed).

Ironically, it was my desire to build smarter legal programs that led me away from law, and back to the problem of agency design—this time, from the perspective of artificial intelligence. In August of 2015, I took Andrew Ng's machine learning course on Coursera. My initial goal was to learn enough natural language processing to automate certain legal work. But four lectures in (the first neural network lecture), something changed—my long-held belief that it was impossible to create intelligence in my lifetime was challenged. This moment triggered my current journey into artificial intelligence. I left law, picked up several texts on math and machine learning for

[1] All papers available at silviupitis.com/papers.

self study, bought a graphics card, enrolled in Georgia Tech's online master's of CS program, and started an academic blog (r2rt.com).

I wasn't sure where this journey would take me, but I knew I needed to learn more. I experimented eagerly, implementing certain results and rediscovering others (e.g., I experimented with what I called "inverting a neural network" before learning what an autoencoder was, and I arrived at the idea of real-time recurrent learning [6] whilst studying backpropagation through time). After two years of study and experimentation, I have more questions than ever and a whole lot left to learn. My attention has shifted away from deep learning per se, toward exploiting it for purposes of artificial agency and bridging the gap between explicit verbal reasoning (logic) and implicit understanding based on data (machine learning).

The first order of business, in my mind, is to continue improving our current models and expanding their range of human-like capabilities. My most recent project on "source traces" endows reinforcement learning agents with a causal model of the environment—a backward-looking take on Dayan's successor representation [1]. I will be presenting my first conference paper on this topic at AAAI 2018 [4]. The main contributions of this paper are a novel algorithm for policy valuation, a proof of its convergence, a novel algorithm for learning the fundamental matrix of a Markov chain, and an exploration of several variations on those algorithms. The eventual aim of this work is to gift agents with the ability to reason backwards, generate hypotheticals, and answer questions like "What might cause X?" Separately, I am working toward developing an interpretable mechanism for reasoning over multiple representations in the reinforcement learning context, on which I will presenting a preliminary abstract at the NIPS 2017 Workshop on Hierarchical Reinforcement Learning [3].

A PhD at the University of Toronto will give me the time, resources and community I need to continue building on my ideas and to fully develop as an artificial intelligence and machine learning researcher. U of T is a global leader for artificial intelligence and offers a wealth of experienced researchers whom I would be honored to learn from and collaborate with. Finally, having grown up in Toronto, it would be an immense privilege to return home for my graduate studies.

References

[1] Dayan, P. 1993. Improving generalization for temporal difference learning: The successor representation. *Neural Computation.* 5, 4 (1993), 613–624.

[2] Pitis, S. 2017. Methods for retrieving alternative contract language using a prototype. *ICAIL '17: Sixteenth International Conference on Law and Artificial Intelligence* (2017).

[3] Pitis, S. 2017. Reasoning for reinforcement learning. *NIPS Hierarchical Reinforcement Learning Workshop.* (2017).

[4] Pitis, S. 2018. Source traces for temporal difference learning. *The Thirty-second AAAI Conference on Artificial Intelligence (AAAI-18)* (2018).

[5] Sutton, R.S., Modayil, J., Delp, M., Degris, T., Pilarski, P.M., White, A. and Precup, D. 2011. Horde: A scalable real-time architecture for learning knowledge from unsupervised sensorimotor interaction. *The 10th International Conference on Autonomous Agents and Multiagent Systems—Volume 2* (2011), 761–768.

[6] Williams, R.J. and Zipser, D. 1989. A learning algorithm for continually running fully recurrent neural networks. *Neural computation.* 1, 2 (1989), 270–280.