Aditya Saraf (sarafa@cs.uw.edu)

ACADEMIC STATEMENT OF PURPOSE

I want to pursue a PhD in Computer Science, with a focus on **theoretical computer science (TCS)**. My goal is to lead a research career after graduating with my PhD, preferably as a professor. Within theory, my interests are broad – I have research experience within **graph theory, complexity theory, and game theory**. I also have a budding interest in Computational Biology and Computational Geometry. I've conducted research in many areas across and outside of computer science and these diverse experiences have strengthened me as a researcher and sharpened my focus on theory.

I started my undergraduate career at the University of Washington (UW) with a concerted effort to get involved in research. An interest in Tech Policy saw me apply for the Tech Policy Lab at UW, and I accepted a research assistant position under Prof. Emily McReynolds during the summer of my freshman year. The lab's goal is to research the privacy and security implications of new technologies to help policy makers build forward-looking laws. During my time there, I looked at three separate topics:

1. Autonomous vehicles. Hacking a single autonomous vehicle could obviously be deadly. But UW machine learning researchers also constructed a (small) physical sticker that caused some autonomous vehicles to mischaracterize stop signs. Clearly, there must be significant policy work (regulating how safe autonomous vehicle technology must be) *before* autonomous vehicles are widely adopted.
2. IoT devices. The privacy issues with always-on recording devices are apparent. And security issues are especially problematic, since compromised IoT devices can form botnets capable of very damaging attacks DDOS attacks. Since we can't expect users to strongly secure their devices, there must be regulation on the security protocols these devices employ.
3. Cell-site simulators. These are devices law enforcement uses to simulate a strong cell tower, which causes all nearby cell phones to connect to it. Law enforcement uses this as a man-in-the-middle attack to view the metadata of any cellular communication in a large radius. UW researchers deployed cell-site simulator "detectors" in Uber/Lyft cars in several major cities to demonstrate that cell-site simulator use was prolific. We looked at the privacy issues behind the deployment of these devices, with particular attention to Fourth Amendment rights.

The most significant project I worked on explored how parents and children react to new internet-connected conversational toys, and this work lead me to coauthor Toys that Listen, a paper published in ACM CHI'17. Publishing a paper in my sophomore year gave me early insight into conducting academic research. My biggest takeaway was learning to identify a gap in the literature and then tailoring a research paper to address that gap – this narrow focus is what enabled us to publish. Given the cross-discipline nature of the lab, I saw first-hand how immersion with diverse ideas led to many research projects. This insight, combined with shifting interests spurred by my CS coursework, led me to focus on Natural Language Processing (NLP) during my junior year.

My initial interest in NLP was born from formal language theory. I took a graduate Linguistics course which covered a "slightly" context-sensitive grammar for English syntax. I loved the theoretical approach to language modeling, but I also wanted a more computational perspective. As a result, I took my capstone class in NLP. At a high level, my project involved comparing neural and combinatorial document summarizers; improving one type of model with ideas from the other category. To do this, we reimplemented and modified state of the art neural models and designed various combinatorial models. This research experience was much more open-ended and independent than what I experienced at the TPL. We had to submit topic proposals, convince our peers of the significance of our work, write a research paper, and create a presentation for our peers and the NLP faculty. This is essentially a microcosm of grad school, so I believe my capstone is a good indicator of grad school success.

My initial interest in NLP was born from formal language theory – but I quickly discovered that the field doesn't currently focus on theoretical foundations. My experience in NLP thus helped me realize that I primarily enjoy the theoretical aspects of Computer Science. This motivated me to apply to CAAR, a summer research

program in TCS organized by Bill Gasarch at UMD. There, I worked closely with Aarthi Sundaram by exploring and fleshing out a framework to study *hidden* constraint satisfaction problems. In this model, an oracle keeps an internal ordered list of constraints and reveals only the *index* of the lexicographically first violated constraint when given a proposed solution. We applied this framework to graphs: we wanted to see if it's harder to find monotone graph properties when a graph is hidden via this "lex-first" oracle. We proved that the hidden setting introduced only polynomial overhead if we could solve the (unhidden) *certificate extension* problem efficiently. Given a graph property, the certificate extension problem asks whether a set of edges can be extended to a valid certificate. As an example, spanning trees certify connectivity, and greedy algorithms can check if a given set of edges can be extended to a spanning tree for a graph (all these notions are formalized in our paper). Having left the hidden setting, we used tools from combinatorics, such as matroids, to proceed. We found that monotone graph properties which have a "matroidal representation" are efficiently extensible – and thus, remain efficient in the hidden context. While this project is still ongoing, we plan to submit to conferences by early 2019.

CAAR also provided me with a broad exposure to theory through weekly theory talks. We discussed current research in Ramsey Theory, Computational Geometry, Cryptography, and Game Theory (among other areas). This program solidified my desire to go to graduate school – doing research fulltime allowed me to get engrossed in one topic, an experience that's not possible while taking a significant course load. I also made good use of my time to network with faculty at UMD. Currently, I'm working with Prof. John Dickerson at UMD to model altruism in dynamic kidney exchanges. Kidney exchanges are systems wherein patients who need a kidney can enroll with an incompatible (but willing) donor, and instead match with another cross-compatible pair. Altruistic donors want to donate a kidney to anyone they are compatible with, without requiring a kidney in return. The productivity benefits of altruism have been well studied in both theory and practice, so researchers want to develop schemes to encourage altruism. Our goal is to develop a theoretical framework for comparing and evaluating such incentive schemes; right now, we're experimenting with mixed Nash equilibria. We model kidney exchange as an n-player game and find a mixed-strategy Nash equilibrium for donating – we interpret this equilibrium as the expected proportion of the population that donates. The details of the game change for different incentive schemes, which then lead to different equilibria. Future directions include applying mechanism design to these games in order to create an optimal incentive scheme. In many ways this project demonstrates the culmination of my research abilities – I chose the topic and project goal, but I'm also working collaboratively with other students and John. Furthermore, I'm mentoring two sophomores interested in TCS by using this project to introduce them to research.

In addition to my CS research experiences, I also performed independent research in philosophy. This work came out of a graduate course in Epistemology (the philosophy of knowledge) under Prof. Conor Mayo-Wilson. My work in philosophy directly improved my academic writing, as philosophical writing emphasizes clarity and conciseness. My own paper in Epistemology explores how agents should use logic as a norm for reasoning, and I'm currently polishing it to submit to journals. I'm also currently working on a project that attempts to weaken the assumptions behind various statistical principles. These principles often require agents to assign degrees of belief to every event in the space and have arbitrarily precise degrees of belief; both assumptions seem impossible to fulfill. We're trying to prove that such principles follow from weaker notions of rationality, such as preference ranking. Though the methods are different, the scope and goals of the project are very similar to papers in game theory. Overall, my work in philosophy complements my work in TCS.

So far, I've described my past accomplishments and how they prepared me for graduate study. Now, I'll look to the future by exploring my motivations for a Ph.D in CS Theory. When I analyzed the complexity of various classes of graph problems, it felt amazing to know that I was creating new knowledge about fundamental properties of graphs. Complexity in general feels like this – we analyze complexity within mathematical models, but the properties proven seem fundamental to the parts of the world we were modelling. Speaking more practically, I also enjoyed applying combinatorics to graph properties. By modelling graph problems with matroids, we were able to use translate well studied tools like matroid intersection to hidden graph properties. Connecting different domains in this way is *extremely* satisfying! Beyond complexity, I enjoy game theory because its concepts seem applicable in many contexts. Many real-world applications are reducible to agent

behavior, as the concept of an "agent" is extremely abstract. I enjoy epistemology for similar reasons – I like answering normative questions about what agents ought to do (and looking at fundamental principles of rationality). Lastly, my interest in synthetic biology comes from a fascination with alternative conceptions of computation – for instance, I've recently been studying a research project at UW and Microsoft Research that is trying to use DNA for long term data storage.

My research experiences in NLP and Epistemology demonstrate my ability to create a research question and dedicate myself to answering it. On the other hand, my experiences at the TPL and through CAAR speak to my ability and desire to form new, productive collaborations with a variety of faculty members. These experiences prove that I can succeed in a PhD program. And in order to achieve my goals, I'm pursuing a PhD in CS theory.