Personal, Relevant Background and Future Goals Statement

I was a writer long before I was a scientist. But it was science that attracted me in high school, took me through an undergraduate degree in mechanical engineering at MIT, and led me into tech-heavy startups. Ultimately I want to marry my existence as a scientist to my existence as a writer. This is what will provide me with the skills, insights and motivation to do important and novel research in natural language processing.

I want to build emotional models of language using a poetry corpus. I believe many roadblocks in sentiment analysis can be overcome if approached from a different perspective. Poetry differs from most natural language in that its primary purpose is communicating emotion, not facts. Most sentiment analysis focuses on emotion as a logical reaction to the facts. Drawing instead from poetry I can develop emotional models that rely on additional features, many of which lie in useful interpretation of figurative language.

A Ph.D in computer science with a focus on interpreting the literary arts will not only allow me to develop the highly technical machine learning skills necessary to build robust language interpretation systems but will also allow me to bring my interests and previous experiences together in a single project. Robust language interpretation systems are the basis for many technology fields that are struggling to meet the needs of society at the level expected, for example speech recognition is useful but not yet robust enough to replace most computer interfaces and sentiment analysis struggles with common language techniques such as metaphor and sarcasm.

As an undergraduate student at MIT I did extensive work with Prof. Hosoi and her graduate student Nadia Cheng on jamming, a technique of applying a vacuum to granular matter which turns a semi-fluid structure into a solid one. I proposed using grains as test samples for understanding how the shape of the granular matter affects the strength of the resulting solid structure. The characteristic length of most grains is within the order of magnitude at which shape properties dominate their shear strength, as opposed to texture or roughness. Their easy availability and multitude of shapes allowed us to test hypotheses around correlations between shape properties and strength. I demonstrated that the 2D shape metrics circularity and polydispersivity were correlated with the strength of the resulting solid asymptotically, an important discovery for those creating robotic structures using this technique. Together with Prof. Hosoi and Nadia Cheng, I wrote a paper based on my discovery, though we never submitted it for review. I did present the results of the paper at the March American Physical Society meeting in 2012.

I also contributed to the design of a highly articulated and underactuated robotic manipulator based on the principle of jamming. I developed a novel locomotive inchworm which used springs as an exoskeleton-style returning force; this concept was then used in the larger manipulator Nadia Cheng designed and presented at the 2012 IEEE International Conference on Robotics and Automation.

These experiences sparked my interest in research. I loved the excitement and thoroughness

of evaluating novel models of the world. Though the topic of this project is very different from language interpretation, it is this experience that gives me the confidence that I will enjoy and succeed in a Ph.D program. For this work I was awarded the Carl G. Sontheimer Prize for Excellence in Innovation and Creativity, given to a single student in the mechanical engineering department at MIT each year.

After graduation I worked as a product developer on the Mimo Baby Monitor which measures the respiration and movement of an infant for the purpose of sending alerts to the caregiver. Developing algorithms for nervous first-time parents requires a high level of accuracy in the detection of events, though it is nearly impossible to run standardized tests with infants. I addressed the main complaint from users, that the sensor did not always accurately detect when it was placed on the child, by running in-house tests on the sensor with an electrically accurate baby model built in conjunction with an electrical engineer. This allowed me to properly threshold our algorithms and resulted in a drop of this complaint from being the most frequent to being so rare that we stopped tracking it.

While working on the Mimo Baby Monitor I also investigated the sleep cycling of infants such that we could inform parents as to the health of their infants' sleep. I based my work on sleep research from the 1980s which indicated that sleep cycling of infants could be detected at about an 80% accuracy using only respiration data. My algorithms detected a night-time period of monitor usage and then ran an analysis to determine the sleep cycling. The results of the analysis were displayed to users as part of a 'recap' of their child's night. This was my first exposure to the complications of interpreting unstructured data, of building on the research of others, and the power data has to provide useful insights. This drove my desire to dig deeper in complex data analysis techniques. I began to read about artificial intelligence techniques. I learned the programming language R. I took Prof. Ng's online introductory course to machine learning. I looked for a job that would allow me to pursue this more sophisticated data analysis.

In my current position as the Research and Development Lead at Soofa, one of the few companies providing smart cities solutions, I led our engineering team to create a pedestrian scanning sensor for our connected park benches. The sensor passively scans for the probe requests that wifi-enabled devices broadcast. Part of this work involved writing firmware for our microcontroller in C, as well as writing firmware for the sensor in Wiring (Arduino's C-based language). Once the sensors were working I was able to begin analyzing the data we were collecting from various cities and parks across the country. This is what I continue to work on today.

In addition to creating simple but utile graphics for clients that demonstrate our ability to detect baseline activity profiles and anomalous events, I work on more sophisticated analysis that attempts to understand the fabric of a community. One area of exploration is finding social linkages between devices by analyzing the name of the Access Points they attempt to connect to. Based on experimental work done by Prof. Cunche at the University of Lyon, I have developed a metric that links devices by measuring the number infrequent Access Point names they share. Another area of exploration is building profiles of types of people who pass by the sensor. I use time-based features of unique pedestrians to build clusters of pedestrians that are seen around the same times each day. After the analysis I interpret the clusters as

types of people, for instance commuters (seen twice a day in the morning and evening) or lunch-goers (seen ones a day in the middle of the day).

All the while I have had an ongoing relationship with poetry and creative prose. As an undergraduate I won several awards for both my own poetry and the poetry I collected for a chapbook entitled 'Poems for MIT Students.' The chapbook was an experiment in writing for a narrow audience, a reaction against the need to prepare your reader for certain references. By specifying that the poems were for other MIT students, the poets were allowed to use language that otherwise would have created barriers to entry for the majority of readers. I raised money for the publication of the chapbook and the cash prizes presented to the selected poets from a variety of places at MIT: the humanities department, the writing department, the art scholars group, as well as the undergraduate association.

In addition, I took several poetry workshops at MIT which introduced me to the techniques of critiquing creative writing in a communal setting. Much of my experience in these workshops has informed how I think about language. Close readings of texts is a dramatically different way to interact with language and highlights the richness of language at the level of word choice; it demonstrates that reactions at the level of an entire text is dictated by a thousand tiny choices by the writer. In poetry, writers are aware (if not down right tortured!) by these tiny choices, though in most texts these choices are subconsciously enacted by the writer.

After graduation I continued this pursuit with poetry publications in a literary journal and a newspaper, a chapbook sold at Grolier's Poetry Shop, along with the publication of a travel essay in Harvard Bookstore's summer travel anthology. It is my interest and experience in creative writing that allows me to bring a unique sensitivity to natural language processing and robust intelligence. Combining experience in the creative arts with machine learning will greatly advance how machines understand and interact with humans above the basic syntactical or topical level.

The reason I am returning to graduate school after three years of technical work in the private-sector is twofold: One, I believe the time allowed for intellectually curious work is too limited by the need to build products that sell well. As I began to talk to professors at MIT and Berkeley about my ideas I loved the way they challenged me to think about my ideas from orthogonal directions. These kinds of conversations are rare in industry where the focus is on the delivery of a feature or product. Two, my recent positions have in no way incorporated my passion for words, which up until this point I have directed solely towards creative writing. However, work in the private-sector, specifically small companies, has shown me how to execute projects on-time and at a high quality. Customers and clients don't care about my technical difficulties and small companies don't have all the resources necessary. All my work has been collaborative but self-driven. I am the kind of person who makes a plan and gets it done.

My experience in the arts will give me the unique perspective to generate exciting ideas about how to further work in sentiment analysis and natural language processing. My experience in industry has provided me with the skills and drive to deliver on them.