Statement of Purpose

Jeewon Hwang Carnegie Mellon University UserID: krjw92@kaist.ac.kr

'Okay. Let's try one last time.'

One evening in the summer of 2013, I was sitting alone in Professor Hensin Tsao's laboratory at Massachusetts General Hospital (MGH). I was there for a Harvard-MIT Health Sciences and Technology (HST) summer internship. I was debugging my code, whose purpose was to make a network graph of inhibitors of melanoma, a type of skin cancer, so that we could find common properties of the most critical inhibitors that suppress gene expression of the cancer. But I had been stuck with an error, and, after five days, I still couldn't find even a single clue why the error was occurring. I desperately wanted to contribute to the treatment of melanoma before returning to Korea, but it was less than a week before the end of internship and it seemed hopeless. 'Okay, let's try one last time', I told myself. I retraced the error as if it were the first time debugging. In a page that I believed to be free of any logical faults, there was the error! I fixed it and pressed the 'run' button. And then, without showing any error logs in the window, the graph finally appeared.

My original motivation for this research stems from my childhood. When I was 8, my mother was diagnosed with breast cancer. This dreadful shock made me want to understand the mechanisms of the disease so one day I could save people, including my mother, from suffering. To this end, I immersed myself in chemistry and biology at my university, KAIST. I spent all my passion and energy on classes and ranked first in almost every course I took. In addition to learning the course materials, I wanted to know how those theories could be applied to medical research, so I visited Dr. Byung-Kook Kim in the biological science department at KAIST. I participated in research analyzing beta-amyloid protein, which appears in neuro-degenerative diseases, such as Alzheimer's disease. Using a circular-dichroism spectroscope and employing the Van't Hoff equation, we could prove that the composition ratio of the protein's secondary structure determines the stability of the protein in terms of enthalpy/entropy. Through this research, I learned that an abstract concept 'unstable' can be quantified into a numerical value. Furthermore, I received a Best Paper Award in the National Research & Education competition.

One day in the library, I came across an article from a scientific journal. It described an algorithm for DNA pattern-matching, which was used for finding gene mutations of maturity-onset diabetes. Prior to the article, I had no idea that computer algorithms could be applied to biological research. Reading the article, I was so fascinated by how computational approaches could give answers to understand the disease. A few months later, I took an 'introduction to programming' course and I found it very exciting. When I entered some logical commands, the computer understood them and returned what I asked. The entire procedure seemed magical! Afterwards, I was strongly motivated to major in computer science and start my journey into the interdisciplinary field that the article had introduced.

In order to broaden my perspective, I first visited KAIST's Bioinformatics and Computational Biology Lab. I participated in a project on drug combinations, with the aim of predicting whether certain combinations of two distinct drugs would have synergistic effects, based on molecular, pharmacological data. We tried to develop an algorithm that models a Protein-Protein-Interaction (PPI) database into a network graph and finds linked paths within it, considering each protein as a vertex and each interaction as an edge. We first developed a recursive algorithm but we faced a problem of insufficient memory. To reduce redundant function calls, I modified the algorithm to use dynamic programming and finally we were able to complete the modeling without the stack

overflowing. From this project, I realized that developing an efficient algorithm is imperative in leading research.
A series of research experiences and my desire finally led me to the Harvard-MIT HST internship. Under the guidance of Professor Hensin Tsao, a faculty member of Harvard Medical School and an expert in cutaneous melanoma, I developed an algorithm that depicted a network graph of inhibitors using existing data from the PubMed database and experimental data about the viability of tumor cells when each inhibitor is used. It allowed us to see common properties of the most critical inhibitors for BRAF and NRAS mutated cells. Once again I learned that developing efficient algorithms is important in research. I applied breadth first search and dynamic programming to reduce the memory usage and increase performance. After constructing the network graph for weeks, we discovered that the BRAF mutated cells are more susceptible to GPR (General Protein Regulator). It was wonderful that I could contribute to research on the treatment of melanoma.

Currently, I am actively participating in a research project that aims to efficiently find 'attractors' in biological systems, under the guidance of Professor Insik Shin in the Computer Science department at KAIST. Attractors are a set of stable states when a biological system is represented in a Boolean network form. Such stable states often correspond to meaningful biological implications, such as cell phenotypes. Thus it is worthwhile to develop algorithms that detect attractors efficiently. I contributed to analysis of the time complexity of our newly developed algorithms, rectifying the proof of the algorithm, and writing a paper. While conducting a series of work, a piece of our work was submitted to *Bioinformatics* entitled "An Efficient Singleton Attractor Detection Algorithm for Large-scale Boolean Network" (C. Hong, **J. Hwang**, K. Cho, I. Shin, 2014, under review).

With my research experience thus far, I have realized that various mathematical models and algorithms are necessary for further interdisciplinary research aimed at understanding biological systems completely. To this end, I will pursue graduate study in computer science in order to strengthen my theoretical background in mathematical models and algorithms. In that reasons, Carnegie Mellon's graduate program in Computer Science is perfect choice for me, both for its academic excellence and various research projects that align with my background. The research on protein 3D structure prediction in Professor Jaime Carbonell's lab attracts my interest. In addition, I believe that my research experience on attractors detection algorithms will be a good match with Professor Christopher Langmead's lab, in the studies of dynamics of biological systems. After extending my knowledge in mathematics and computer science, I hope to continue my career in the 'Ray and Stephanie Lane Center for Computational Biology'.

I know it will be a challenging environment along the way to one's own research, and the one who is resilient to those difficulties will be the expected candidate of your program. I already had experiences of overcoming such difficulties during my undergraduate years and in that sense I am confident I can excel in such challenging environment again. When I first started to learn computer science in KAIST, it was extremely difficult to catch up with the class. In systems programming, the first course I took in computer science, I was at the bottom of my class for the midterm exam since I had never experienced programming in C before. It was desperate, but I trusted in myself that it would all work out and I kept doing my best. Eventually, I achieved a high major GPA (4.01 out of 4.3) by the time I graduate and I received scholarships from the Korean President, Google, and Samsung. In this regard, I believe that my academic achievements demonstrate my potential, and I

am certain that I will be an outstanding addition to your program, as much as CMU will be to me.