



Department of Computer Science

December 13, 2004

Professor Rob de Ruyter van Steveninck
Faculty Search Committee
Biocomplexity
Indiana University
Department of Physics
Swain Hall West 117
Bloomington, IN 47405-7105

Dear Professor van Steveninck:

This is a letter of recommendation in support of Dr. Mingzhou (Joe) Song, an applicant for a faculty position in the Biocomplexity at Indiana University.

I know Joe since 1997 when he came to University of Washington, Seattle, as a Ph.D. student, to study under Professor Robert Haralick. During Joe's Ph.D. study, I was a professor at Seattle University and I was a graduate affiliate professor at University of Washington. I conducted my research at University of Washington in the same laboratory as Joe did his Ph.D. work. For almost four years, I saw Joe on a daily basis until fall 2000 when I took the chair position in the Computer Science Department at Queens College. Upon his graduation in 2002, Joe accepted a faculty position, as an assistant professor, in my department at Queens College.

Before I give my recommendation on the qualifications of Joe, I should mention that it would be a great loss for my department if Joe leaves. However, I also know and understand the reason that Joe wants to find a better position elsewhere.

There is a Computer Science Ph.D. program in the Graduate Center at the City University of New York (CUNY), where Queens College is under its umbrella. However, my department at Queens College does not have a Ph.D. degree program on campus. In addition, Queens College requires a six-course teaching load per year. With no teaching assistant and very limited resources provided, there is little time left after teaching for most faculty members to do research. Under such circumstances, Joe has managed to supervise a couple of Ph.D. students at Graduate Center. However, Joe was trained as a researcher from University of Washington and would like to find more resources to pursue his research goals.

Joe discussed his plan of leaving the current position with me. I honestly understand his wish to go to a department that offers a greater opportunity for research.

Now I shall comment on Joe's qualifications for the position in your department. Judging from my experience with Joe for the past seven years, I am fully convinced that Joe has been a high quality researcher and an effective teacher.

On his Ph.D. committee, I was able to closely keep up with Joe's research progress. During his

doctoral study at University of Washington, Joe worked in two areas – statistical pattern recognition and 3-D medical imaging. He addressed the important optimal quantization problem, which determines the optimal grid partition of space and the probability density value of each grid cell based on a criterion involving likelihood, entropy and classification accuracy. Joe has successfully designed an algorithm to obtain an optimal solution in one dimension by a dynamic programming approach. Joe found that the dynamic programming algorithm is not extendable to multidimensional space. Thus, he developed another approach that combines a global optimization by genetic algorithms and a local optimization to find a reasonable grid for the space based on the given data set. The work on optimal quantization relied heavily on his computational and statistical expertise. I believe that this result by itself qualifies Joe's research as immensely important for the research community.

The 3-D image reconstruction project was collaboration with the School of Medicine at University of Washington. The goal of this project was to reconstruct 3-D geometric models for the left ventricle using images acquired by freehand ultrasound imaging machines. It was the low quality of ultrasound images that made the project especially challenging. A traditional computer vision approach uses a two-level strategy, i.e., low and high levels. However, this approach fails in low quality images most of the time. Inspired by observing the manual left-ventricle boundary tracing practice of sonographers, Joe came up with an approach to combine the two separate level solutions into one overall optimization step under the Bayesian framework. With this optimal quantization method, Joe constructed a probability model to capture the features of image pixels. In this work, he represented the high-level shape information of left ventricle using a convex combination of exemplar shapes. The low- and high-levels are linked together by a pixel class predication probability model. The probability models are estimated by the expectation maximization algorithm considering low- and high-level simultaneously. The final left ventricle model is obtained by maximizing its posterior probability given the images, again in a combined low- and high-level scheme. This approach was exceptionally novel. As the result, the method has obtained the best left-ventricle modeling results ever reported in literature. The work was published in IEEE Transactions on Medical Imaging, the most cited journal in the field of medical imaging.

In terms of his research at Queens College, Joe's energetic efforts have resulted in a continuous stream of top journal and leading conference proceeding publications. After Joe came to Queens College, he followed his Ph.D. research of optimal quantization. Joe has applied it in recombination rate estimation for human genome and spike sorting for neuronal signals. The result of recombination rate estimation was published in the leading conference IEEE Computational Systems Biology Conference in 2003 at Stanford, CA. Joe presented the spike-sorting work at the Joint Statistical Meetings in 2003 in San Francisco.

Joe is an independent researcher who is able to identify and establish his own area of research. As his colleague at Queens College, I was able to observe Joe's exceptional abilities to develop his own research theme on data stream clustering algorithms. A data stream clustering algorithm maintains clusters dynamically without keeping all historical data in memory. Most previous work assumed spherical clusters and adopted distance-based approach. His idea is to use density-based

approach and to maintain sufficient statistics for the clusters from historical data. Densities of newly arrived data are merged with current clusters by using the sufficient statistics. Joe's initial result in this project has just been accepted by the SPIE Conference on Intelligent Computing for presentation in March 2005 in Orlando, FL.

Joe's career so far shows that he is able to interconnect his research in both bioscience and computer science/engineering disciplines. Indeed, his cooperative research activities have been greatly complemented by other researchers in bioscience. He has been working with biomedical researchers on several systems biology and computational neuroscience projects such as computing gene regulatory network, modeling repetitive genomic elements and their evolution, neuronal signal classification and clustering, and microscopic image analysis for brain cells.

Joe's research performance has been remarkable, regardless of the heavy teaching load and limited research support at Queens College. With his excellent research proposals, Joe won several CUNY research grants. These grants allow him to pursue and implement his research plans. Joe also continuously seeks external fundings. Recently, he assisted, for Queens College, in obtaining an external award from the Howard Hughes Medical Institute for Undergraduate Science Research Program. This program provides support for several postdoctoral fellows and new undergraduate curricular development in the interdisciplinary area that crosses biology and quantitative sciences such as bioinformatics.

In terms of teaching, Joe is an excellent and devoted teacher. He cares about his students. For every lecture he teaches, Joe prepares thoroughly ahead of time. He poses his lecture notes on his web site for his students. He listens to his students when they come to him. He enjoys seeing the steady progress made by his students. I have sit in several of Joe's classes to observe his teaching. He remembers student first names and encourages every student to participate actively in class. His presentations are clear and skillful. He makes difficult topics easy to understand for students.

At Queens College, Joe has developed a course on Introduction to Bioinformatics, the first such course at Queens College. Other courses he has taught included Algorithm Design and Analysis, Artificial Intelligence, Data Communications, Research Practicum (in bioinformatics) and Software Practicum (in microscopic image analysis), as well as the Topics in Computational Biology at CUNY Graduate Center. Another important educational agenda for Joe is to work with undergraduate students to develop their research skills. He mentored several undergraduate students for research. He had a paper with an undergraduate student in Joint Statistical Meeting in San Francisco in 2003 on a neuronal spike-sorting algorithm. He has also been a member on behalf of Department of Computer Science on a college-wide planning committee for a new Master's program in bioinformatics.

In terms of communication skills, Joe has been studying and working in U.S. for the past eight years. His spoken and written English are excellent. He communicates easily with professionals in both natural and physical sciences. He believes in the value of effective communication and makes every effort to strengthen his ability to communicate with others.

Joe's personality is very pleasant and courteous. He gets along with colleagues and blends well

into the department.

Joe is an excellent researcher with full potential. He is also an exceptional educator. Without any reservation, I give my strongest recommendation for Mingzhou (Joe) Song. I would expect him to be a very productive member in your department.

If you have any questions, please do not hesitate to contact me.

Sincerely yours,



Professor Tsaiyun Ihsin Phillips, Ph.D.
Department of Computer Science
Queens College of CUNY
65-30 Kissena Blvd
Flushing, NY 11367
Ph: (718) 997-3524 (office)
Ph: (516) 374-1143 (home)
Email: yun@cs.qc.edu