AC 2010-2389: COMPUTATIONAL SCIENCE AND ENGINEERING EDUCATION, RESEARCH AND TRAINING AT A HBCU - EXPERIENCES AND OUTCOMES

Ram Mohan, North Carolina A & T State University

Ajit Kelkar, North Carolina A&T State University

Narayanaswamy Radhakrishnan, North Carolina A&T State University

Computational Science and Engineering Education, Research and Training at a HBCU – Experiences and Outcomes

Abstract

Computational science and engineering (CSE) and high performance computing (HPC) have now become an integral part of several engineering and science disciplines. Still the number of students from under-represented universities and minority institutions who are involved and educated in these fields is very minimal. The interdisciplinary computational science and engineering graduate program at our University (a major historically black college and university) was initiated in 2005 and has now graduated more than 25 students (including minorities and women, and in particular, African Americans) from various disciplines of science and engineering. This graduate program is the first such program at a historically black college and university (HBCU) in the nation, and is highly interdisciplinary, offering a curriculum that combines the stronger domain knowledge of several areas of science and engineering with the CSE core areas of applied numerical analysis, software data structures, parallel and high performance computing and data visualization. The interdisciplinary nature of the program and the presence of students from various diversified backgrounds pose significant challenges in the core computational courses. The details of the CSE educational program, experiences and outcomes from this graduate program are discussed in this paper. Representative case studies of the experiences and outcomes of the students from the program over the years are highlighted.

A complementary outreach aspect of the CSE educational program at our university is the annual workshop on CSE and HPC for faculty members from other minority serving institutions across US and Puerto Rico. This educating the educator workshop has benefited the participating faculty, allowing them to not only get trained and exposed in these areas but also act as catalysts to propagate their knowledge to their students, serving a larger minority population. Our experiences and outcomes from the past few years and their impact on the CSE education and training of under-represented minorities are highlighted.

The positive outcomes of the initial Master's level graduate program are now enabling the establishment of a PhD program in CSE (first among HBCU in the nation) that is currently pending approval for establishment. Details of the PhD program development and structure are briefly presented.

Introduction

The new paradigm in graduate studies is interdisciplinary programs that meet the technical needs of the current practices in the field and global industry. Modeling and simulation built upon computational science and engineering has now become the third key solution methodology in not only engineering and physical sciences but also in other areas such as biology, economics, and health sciences that are generally considered to be non-computational fields. The interdisciplinary master's degree program in Computational Science and Engineering (CSE) at our University is now more than 4 years old, and provides graduate education in several computational areas and the associated primary field disciplines. The CSE program since its inception has graduated more than 25 students who are currently placed in several major global industries.

Our CSE graduate program offers an interdisciplinary curriculum combining applied mathematics, high performance parallel and scalable computing, scientific modeling and simulation, data visualization, and domain areas such as physical sciences and engineering, life sciences, agricultural and environmental sciences, technology and business. The students enrolled in the program begin with diversified backgrounds (prior undergraduate studies in engineering, physical sciences, life sciences, mathematics, business, etc). However, all students in the program are required to take four core courses relevant to CSE. These are: comprehensive numerical analysis; scientific visualization; applied probability and statistics; parallel programming and data structures. The preparation level for the diversified group of students in these courses depends on their undergraduate major. This poses significant challenges to graduate faculty teaching these courses and mentoring these students with diversified backgrounds.

The outreach and popularity of the new areas of computational science and engineering to other minority serving institutions and population can only be enhanced if the other institutions and their students are exposed to them. This requires that the faculty members at the minority serving institutions themselves are exposed to the area. A complementary outreach aspect of the CSE educational program at our university is the annual workshop on CSE and HPC for faculty members from other minority serving institutions across US and Puerto Rico. This educating the educator workshop has benefited the participating faculty, allowing them to not only get trained and exposed in these areas but also act as catalysts to propagate their knowledge to their students, serving a larger minority population.

This paper is organized as follows. The interdisciplinary graduate master's program and its structure at our institution are discussed first. This is followed by brief discussions of our experiences and observations in teaching and mentoring the diversified group of students with varying incoming preparation levels and undergraduate degree background in two of the core courses. An overview of the faculty training program is briefly presented next. The positive outcomes of the initial Master's level graduate program are now enabling the establishment of a PhD program in CSE (first among HBCU in the nation) that has been approved for establishment and approval from fall 2010. Brief details of the CSE PhD program curriculum are presented and discussed.

Computational Science and Engineering Program at Our University – Overview

Our CSE master's program has three underlying tracks with a focus on computational science, but distinguishes among the domain areas of specialization. The three tracks have a common curriculum through four required core courses and accounts for breadth and the variations in computational science field requirements across the several domains. These core courses are required to be taken by all the students in the program. Thus the students in these core courses are a diversified group with a different preparation level depending on their background and undergraduate major. The three major underlying tracks with a common core curriculum required for all the students are briefly presented next.

Computational Science and Engineering

This track is designed primarily for students with undergraduate degrees in engineering, physics, mathematics, and computer science who will are trained to develop problem-solving methodologies and computational tools as well as interdisciplinary technical expertise in CSE for solving challenging problems in physical science, engineering, applied mathematics or computer science. The curriculum emphasizes computational sciences and engineering along with training in the domain areas. The goal of this track is to produce scientists and engineers with focus, training and application in computational sciences, scalable computing, and physics-based modeling and simulations, and with expertise in the application of computational techniques and principles in their primary domain areas. Based on their undergraduate degrees, the students in this track have had an increased level of prior training, courses and exposure to mathematics including areas such as numerical analysis and in some cases to high level programming languages.

Computational Sciences

This track is designed primarily for students with undergraduate degrees in chemistry, biology, business, and agricultural sciences who are trained to apply or extend computational tools and methods as well as data acquisition, processing and visualization techniques to study computationally intensive problems in life sciences, agricultural and environmental sciences, and business and economics. This track primarily includes domain areas with less training in mathematics including numerical analysis, programming languages.

Computational Technology

This track is designed primarily for students with undergraduate degrees in technology disciplines with focus on computational science and engineering. These technology disciplines currently include computation technology, computer numerical control machining, remote sensing, and GIS/GPS data analysis, with additional potential disciplines in the future. The goal of this track is to produce technologists with a focus and training in computational sciences, and in their primary technology domain area.

Educational Objectives of our CSE Graduate Program

The educational objectives of our CSE graduate program are as follows:

- Educate and graduate students with a mastery of high performance computer programming tools as well as data processing, data acquisition, and data analysis techniques.
- Educate and train students in computational modeling, simulation and visualization.
- Promote learning across the boundaries between engineering, science, technology and business.
- Relate acquired computational science and engineering knowledge and skills to specific application fields of engineering, science, technology and business.
- Teach students to develop novel and robust computational methods and tools to solve scientific, engineering, and technological and business problems.
- Produce highly versatile computational scientists, engineers, technologists, or business executives with a good understanding of the connections among various disciplines and of

the capability to interact and collaborate effectively with scientists, engineers, and professional in other fields.

- o Increase the number of graduate professionals available to work in computational science and engineering.

Course Curriculum for CSE

CSE has emerged as a powerful and indispensable method to analyze a variety of problems in research, production and process development, design and manufacturing, as well as in domains such as computational chemistry, biology, genomics, business forecasting, economic modeling, etc. Computational modeling and simulation is being accepted as a third methodology in scientific discovery processing and engineering design, complementing the traditional approaches of theory and experiment. Many experiments and investigations that have traditionally been performed in a laboratory or the field are being augmented or replaced by computational modeling and simulation

Even though CSE makes use of the techniques of applied mathematics and computer science for the development of numerical algorithms and computing tools to the study of scientific and engineering problems, as well as other physical, biological and non-scientific areas, it is by no means a subfield or extension of applied mathematics or computer science, nor is it a discipline where a scientist or engineer or domain specialist simply uses a canned code to simulate data and visualize results. "CSE is a legitimate and important academic enterprise," as noted in a comprehensive, report [5] published by the SIAM Workgroup on CSE on Graduate Education in CSE. "Although it includes elements from computer science, applied mathematics, engineering and science, CSE focuses on the integration of knowledge and methodologies from all these disciplines, and as such is a subject which is distinct from any of them." Figure 1 presents the interdisciplinary nature of the CSE program, and reflects the view that besides connecting the sciences, engineering, mathematics, and computer science,

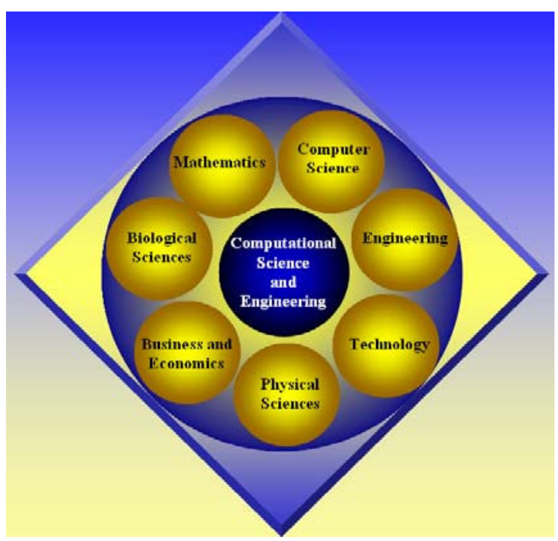

Figure 1: Interdisciplinary Nature of the CSE Program

CSE also has its own core of elements that draws together and bridges all these disciplines. Such a CSE core is made up of a collection of computationally intensive problem-solving methodologies and robust tools, which constitute the building blocks for the study of scientific and engineering problems and non-engineering applications of ever increasing complexity and realism. Based on the above core requirements of CSE education and objectives, our CSE program requires all the CSE students to take four CSE core courses irrespective of their undergraduate background and CSE tracks. These courses provide the foundations of the enabling technologies of the computational sciences for all students irrespective of their background. These core courses are:

CSE 701: Applied Probability and Statistics
CSE 702: Comprehensive Numerical Analysis
CSE 703: Parallel Programming and Data Structures
CSE 704: Computational and Scientific Visualization

These four core courses are taken by all the students in their first three semesters of their graduate study. Thus the students taking these classes have a varied preparation level, background and exposure to some of the pre-requisite concepts that are generally needed in these core courses. The preparation level for the diversified group of students in these courses depends on their undergraduate major. This poses significant challenges to graduate faculty teaching these courses, evaluating, and mentoring these students with diversified backgrounds. Our experiences and observations with the course content and structure, teaching methods, evaluation and student performances in two of these four core courses with diversified graduate students and their mentoring are discussed next.

Experiences and Observations in Teaching CSE Core Courses

This section discusses our experiences and observations in teaching and mentoring the diversified graduate students in the CSE courses. Our experiences and observations from CSE 701 which is also cross listed as a course in Biology are presented next.

Experiences and Observations from CSE 701: Applied Probability and Statistics

Most students taking CSE 701 – Applied Probability and Statistics are new Master's students in their first year of graduate school. It is always wonderful to teach new incoming graduate students who come in focused and on a mission. Several have experienced the real world outside their alma mater and are looking forward to acquiring new skills and degrees that will make them professionally more competitive. However, most students have entered graduate school directly from their undergraduate semesters, armed with Bachelors degrees, but unhappy with the job market and their career prospects.

Just as their motives for being in this class are diverse so are the academic backgrounds of the CSE graduate students. Many students have undergraduate degrees in mathematics or engineering, whereas some have majored in biology, business, physics, and agricultural sciences. Their level of knowledge of statistics is also varied, although, most have had at least some elementary statistical theory. Teaching statistics to students from different backgrounds is not difficult, statistics is, in essence, an applied science and lends itself well to most disciplines. However, dealing with all the different backgrounds and skill levels at the same time can sometimes be a challenge.

The content of this core course in applied probability and statistics is traditional, in the sense that it covers both descriptive and inferential statistics for various types of data that will be useful for different disciplines and a necessary tool for computational scientists. Data transformation techniques often required for statistical modeling of real life data are discussed, along with non-parametric methods for when parametric assumptions are not satisfied. Fundamentals of statistical and research design methods are also covered to train and educate the students about to embark on their graduate theses and projects. Statistics software such as SAS and/or SPSS is used and tailored according to students' skill levels and disciplines. Along with the conventional examples and problems (often based on agricultural research data), analysis of cutting edge research data from several disciplines is discussed. Some discipline research data are: Statistical modeling or designing experiments in composite material research; Experimental design and analytical issues involved in the analysis of gene expression data from Microarrays in computational biology. The objective of these different examples and problems is to provide relevance of the statistical modeling to each student's own background and thus to provide the increased motivation for the students. These multidisciplinary examples also provide an understanding of the cross-disciplinary applications of the principles of statistics and their relevance in several fields. This being the case, this course as taught as a CSE core course is not designed to provide any specific entree to satisfy every appetite, but a smorgasbord of statistical concepts, tools and techniques, that are available and can be made relevant to most types of research studies and subsequent data analysis.

Our observations and experiences indicate that one unique aspect of working with or teaching an interdisciplinary course such as this is that the teaching or mentoring rarely ended with the end of the semester or even with an 'A' grade. The statistical concepts introduced to the students have made them realize that in a research study the time to see a statistician is not when the data has been collected but when planning for the data collection. This was found to be the case in a few case study observations. Examples of such experiences are:

- "John Smith" came to see me this summer when he started planning his study to discuss his research design and its validity and of course the type of data he plans to collect and the appropriate methods to analyze.
- "Jane Jones" knew she would need to use regression models for her analysis, but we had only covered linear regression in class, "and came to discuss if some other model be more appropriate?" - She was willing to research and study some advanced models on her own.
- "James R" was working on an agri-business project for his thesis, it involved a large data set with missing values that required some imputation; He knew SAS was good at handling these types of datasets needed help with debugging some do-loops.

Teaching interdisciplinary students is an invigorating experience for faculty that challenges one to think outside the box of your own field or discipline, and in finding effective ways of teaching these students.

Experiences and Observations from CSE 702: Comprehensive Numerical Analysis

CSE involves the use of computational hardware architecture and associated software to develop numerical algorithms or methods to study scientific, engineering and other application domain problems. Most mathematical models in the physical sciences and engineering as well as in other non-scientific disciplines involve the solution of systems of linear or non-linear equations, differential equations, integral equations or similar type of problems. As such, all graduates in CSE should be educated and have knowledge of the numerical algorithms for the solution of such mathematical models, their computational implementation, the computational complexity and cost of such algorithms, and the computational errors and accuracy of such computational algorithms. The core course of comprehensive numerical analysis was developed to provide a foundational background of these techniques for the diversified backgrounds of the incoming graduate students.

Traditional courses in numerical analysis cover in depth the algorithmic methodologies, their formal mathematical development, errors and accuracy if the course is taught as a formal mathematics course. A similar course in computer science generally focuses on the efficient computational implementation of the step by step numerical algorithms, computing errors, computational complexity of the algorithms and the computational cost. Computational scientists of any domain discipline should have a good understanding of the underlying computational algorithms, limitation of such algorithms, computational implementation and computational cost of the algorithms employed for the solution of the mathematical model equations. The comprehensive numerical analysis course (CSE 702) was developed and formulated to include this understanding of the computational algorithms (with a reduced emphasis on the formal mathematical development of the algorithms) and their computer code implementation, computational cost comparisons of different algorithms that are possible for the solution of the model equations, and the computational caveats (things to look out for that can lead to incorrect results) in such computational algorithms. The CSE 702 course content and student evaluations thus included the following major components:

- Exercises and problems to formally evaluate the understanding of the computational algorithms and the computational cost with simple system of model equations.
- Computational code implementation of the algorithms in a high level programming language to understand the practical aspects of such implementation and understand the caveats of computational algorithms via a larger set of representative problems involved in the numerical solution. Students have the option of computational implementation in any high language they are familiar. These implementations are necessary as an integral part of CSE education as the CSE graduate education is to go beyond the general use of canned analysis software, and understand the development and the associated computational issues of computational algorithms and their implementation.

These two major components required students to develop proficiency not only in the basic analysis that involved the understanding of the algorithm logic and mathematical principles, but also in the programming and code development background.

Student Background in CSE 702

The students in CSE 702 had diverse backgrounds with undergraduate and graduate study in various disciplines. The major grouping of the backgrounds and the associated strengths and weaknesses observed based on our experiences are presented next.

Physical Sciences: This included students with an undergraduate degree in one of the physical science areas of physics and chemistry. Though most of them had at least two courses in calculus, a course in differential equations, and some introduction to linear algebra, many did not have a formal course in numerical analysis. Surprisingly, they also appear to have very limited experience in any high level computer language at our University.

Engineering: This included students with different domains of engineering. The students with the bachelor degree background in industrial engineering were found to have a lesser mathematical background (calculus, basic linear algebra) compared to the other domains of engineering (mechanical, chemical, etc.). The CSE incoming class students also were found lacking in the area of formal programming languages, with many having limited experience in code implementation, compiling and debugging and had only prior experience in canned codes such as MATLAB functions.

Computer Science: The students with an undergraduate background in computer science had a prior background in some formal programming language and an algorithmic mindset. However, they lacked the physical connection associated with the problems.

Mathematics: The students with an undergraduate profile in mathematical sciences had a broader base in advanced calculus and primary numerical analysis but surprisingly lacked the code implementation and high level programming language skills.

Technology: The students with undergraduate background in technology lacked a formal mathematical background beyond the basic courses in mathematics, with several having negligible experiences in code development and implementation.

Business & Economics: The students with this undergraduate background generally lacked the mathematical background (other than preliminary calculus courses) and formal programming background.
This diversity in student backgrounds led to the following experiences and observations and required modifications during the teaching of the course based on the class background. Each batch of students was unique but with the following commonalities observed.

- Most of the students lacked prior computational code implementation, development and implementation. This required additional class handout material to cover these topic areas. Generally the students used either the UNIX based computing systems for their code implementations or the Visual C, Visual C++, Visual Fortran compilers on Windows based machines. Many were not familiar with this aspect of code development, compilation and code execution that are not generally associated with the use of canned computing modules such as MATLAB. Additional class materials and tutorial sessions were required to bring the lagging students up to speed. Interestingly, most students were able to understand the algorithms, but had several issues in the computer code implementation syntax due to the lack of prior code implementation background.

- Generally, the students with a computer science background had a faster turn around time in the completion of the computer implementations due to their formal training in the data structures and computer code languages and syntaxes. These stronger students were also able to provide peer guidance to the weaker students and geared them up quickly. Several students also required extra time for the completion of the computer implementation assignments during the initial assignments with the improving performance as their experience and confidence level increased.

- Students who are weaker in mathematical aspects required some additional assistance, but were able to gear up quickly with focused, individual guidance.

- The motivation level and the willingness to strive and succeed were observed to play a major role in the computer implementation irrespective of prior student background.

All the observations are based on the experiences at our university. The data suggests that the performance is related not only to an individual's prior background but also to their willingness and motivation to strive and succeed. The data based on student population at our university suggests that though it is generally expected for the mathematics trained students to perform well in numerical analysis courses, it was not the case in some instances as it also depends on the student's effort level. Some of the non-science and engineering undergraduate majors out performed the math majors. Our mental note of the observations during the course indicated a higher motivation and willingness to strive and succeed as a factor in such cases.

Remarks and Observations:

The teaching and mentoring of students with diverse backgrounds is definitely challenging and required some out of the box approaches for effective communication and teaching of materials along with additional training in areas of programming, code development, etc. This area was found to be weaker in our experience at our university. It is clear that the incoming students need additional training in code development, other computer operating systems, and high level programming languages. The training in these areas can be achieved through additional short training workshops in the first two weeks of the first semester of the incoming CSE graduate student. Our experiences and observations indicate such short training courses for the students should include the following areas that would benefit the CSE graduate education and training.

- High Performance Computing Operating Systems (Introduction to Unix/Linux)
- Practical Programming in C/C++/Fortran

Our CSE graduate program provided the students with an interdisciplinary education with the core courses contributing to this interdisciplinary education. Generally students have a perception that disciplines and disciplinary methods are distinct and disciplines do not a commonality. The core courses of our CSE program are providing the students with an effective interdisciplinary thought process and bringing out the commonality of applicable techniques across the disciplines. This has resulted in the career placement of graduates in areas that generally would not have been possible based solely on their undergraduate field. Some representative cases in our experience are:

- "Jane Doe" with an undergraduate degree in mathematics, and minimal prior programming experience obtaining a "Computer Scientist" position after graduation with CSE degree.
- "John Smith", an undergraduate in computer science, with minimal engineering experience developing expertise in finite element modeling and large scale, parallel codes in mechanical engineering. These cross-disciplinary understanding is a positive outcome of the interdisciplinary graduate program and mentoring.
- "Mary John" a mathematics undergraduate major with minimal science background working in the area of computational material science.

Annual CSE HPC Minority Faculty Training Program

Computational science and engineering (CSE) and high performance computing (HPC) have now become an integral part of several engineering and science disciplines. Still the number of students from under-represented universities and minority institutions who are involved and exposed to these fields is very minimal. Several high performance computing training programs funded by the National Science Foundation, the Department of the Defense, etc., have traditionally focused on summer training in HPC for minority students from minority serving universities. In spite of several such programs, the number of minority, especially African American students, graduating and entering CSE and HPC professional areas has remained really low. This can only be met by the introduction of these technologies, education and mentoring of the under-represented and minority students at various levels. This could potentially be facilitated by an early introduction of these technology areas from the undergraduate level. It would be best if the undergraduate students are exposed to such areas by their own faculty advisors, teachers and educators at their own institutions. This would permit this exposure to reach a larger segment of the minority student population than are reached by the current HPC summer student programs. In most cases, the students also lose continuity and nurturing after the completion of the summer programs due to lack of local faculty expertise and interest in these areas at most of the minority serving institutions. It is thus imperative that faculty members themselves at these minority and under-represented serving institutions be exposed and trained in the field of high performance computing, and the related technologies that enable the computational modeling and simulations, and the application of the high performance computing techniques, methodologies and paradigms.

Our experiences indicate that the practical high performance computing knowledge expertise and skills are generally learned during their working use and practice by most scientists and engineers. At any minority institution, most of the faculty members have expertise and formal training in their own fields, and in the domain area courses they teach. Some of the faculty members who have graduated a while ago may not have any prior formal training in the area of computational science and engineering, and high performance computing. It is thus essential that the faculty members at these minority serving institutions be exposed and trained in the areas of computational science and engineering, and high performance computing.

In this regard, over the past four years, we have followed a complementary approach by exposing and training the faculty members from several under-represented and minority serving universities in the areas of CSE and HPC through a one-week annual workshop conducted at our university complementing our computational science and engineering graduate program. Our experiences, findings, and data indicate that the workshop participant faculty members from under-represented and minority serving institutions have not had any prior exposure to various aspects of the CSE and HPC techniques and opportunities. This workshop facilitated and provided such training to nearly 60+ faculty members from 40+ different minority serving institutions over the past 4 years.

The anonymous feedbacks from the participants clearly indicate that this educating the educator workshop has enabled the participating faculty members to get trained and exposed in these areas. In addition these faculty members also act as catalysts to propagate their knowledge to their students. This serves a larger minority population and has a potential to provide future work force needs of qualified minorities in the critical technical area of computational science and engineering and high performance computing. The workshop is not only enabling the minority serving institutions (MSI) faculty member improvement and exposure to new areas but is also benefiting the student education and mentoring at minority serving institutions.

CSE PhD Program

The MS graduate program and the annual workshop have been the catalysts behind the establishment of the PhD program in the area, the first such a program among HBCUs in the nation. The program structure follows a similar pattern to the MS program.

Ph.D. Program Courses

Courses for the CSE MS and PhD degree program build upon courses in existing master’s degree programs in the sciences, engineering, mathematics, technology, and business, yet address the goals and objectives of the CSE program. Many of the disciplines and degree programs on campus have in place courses that support the CSE master's degree. For example, numerical linear algebra, numerical PDEs, scientific visualization, distributed and high performance parallel computing, computer organization and scientific programming, data structure, software tools, and computational science and engineering courses already exist in applied mathematics, computer science, physics, biology and mechanical engineering master’s programs. Selected existing courses were modified as core courses and others as interdisciplinary or domain elective courses for the program. A few new courses including several core courses for computational sciences and computational technologies as well as a couple of bridge courses were developed for the program.

Figure 2 presents the overall curriculum structure for CSE PhD program. The PhD program requires 50 credit hours and fits the “T” paradigm of the graduate education providing the depth and breadth required for the next generation of graduates.

Figure 2: CSE PhD Curriculum Structure

Concluding Remarks

Computational science and engineering (CSE) and high performance computing (HPC) have now become an integral part of several engineering and science disciplines. Still the number of students from under-represented universities and minority institutions who are involved and educated in these fields is very minimal. This paper presented our successful CSE graduate program highlighting the growth, experiences and outcomes, difficulties faced, and positive impact on the CSE education and training of under-represented minorities. The positive outcomes of the initial Master’s level graduate program are now enabling the approved establishment of a PhD program in CSE (first among HBCU in the nation) from fall 2010.

Acknowledgments: We thank and appreciate the contributions and the experiences of Dr. Vinaya Kelkar who teaches the CSE 701, Applied Probability and Statistics. Further details of the discussions and experiences from CSE 701 are presented in “Experiences in Teaching and Mentoring Interdisciplinary Graduate Students of Diversified Backgrounds”, AC 2008-2901, ASEE.