REU Site: Interdisciplinary Program in High Performance Computing
Student Activities

The following text was originally excerpted from our NSF proposal for the renewal of the REU Site program in August 2011. We promised in that proposal to make it available on the webpage, because this is the most concrete description of the activities and might be useful to prospective applicants to see.

(a) Overview

The student activities of this REU Site program are delivered in three overlapping phases. As explained in more detail below, the **formal instruction** by the PI/co-PIs, assisted by graduate assistants, in **Phase I** of the program will be explicitly used to establish the project teams, to give them experience with team work and cross training each other, and to create the comfort level necessary for the students to approach the graduate assistants and others for help. We believe that it is most effective to establish this environment first in a formal instructional framework familiar to college students, while building the team work, communication, and other skills needed for the more unstructured **Phase II** of **team work on application problems**. Project examples from the past show the wide range of topics available from project mentors who might be from industry, government agencies, other departments at UMBC, or other universities. Some formal instruction will continue during Phase II on many other aspects of professional importance, such as professional integrity, graduate school preparation, and some advanced topics of high performance computing, in addition to topics directly related to the project work such as presentation skills (both informal and formal) and writing skills (both technical such as LaTeX introduction and conceptual such as the structure of a journal paper and the submission/review process of publishing). The latter skills are needed in **Phase III** of the program on **dissemination**, in which final reports and presentations will be prepared. In practice, Phases II and III naturally overlap, but we identify Phase III explicitly (also to the students!) to make clear that the final outcome of research are presentations and possibly scholarly publications of the work. Throughout all phases, we will provide ample opportunity for informal and semi-formal communications with the PI/co-PIs and graduate assistants. In particular, the purpose of tightly integrating graduate assistants in this framework is to open communication channels to the undergraduate students with people closer to them in career path than only faculty and professional scientists.

(b) Nature of Student Activities

This section details how the principal educational components appear in the three phases and expands this classification by describing how a student participant would experience the timing. This will reveal that Phases I and II purposefully have some overlap, and that Phase III is in reality built into the last 5 weeks of the program. The following detailed description of the three overlapping phases will be available on future program webpages for best possible information for prospective applicants, in addition to the existing information including a detailed schedule that shows day-to-day activities down to the hour.

Phase I - Weeks 1 and 2: The goal of this REU is to involve students *actively*. To this end, Monday of Week 1 starts with a lecture that shows the state-of-the-art computing cluster, how to log in, write C code, make parallel code, and its possible performance. A computer lab immediately following lets the teams repeat these tasks themselves under the watchful eye of a TA to ensure that no team gets stuck and all members are involved. Performing these tasks live gives the teams an opportunity to figure out how to work together and how to jointly reach out to

TAs and faculty for assistance. Each day of Weeks 1 and 2, except Thursdays, consist of 2 lectures and 2 computer labs per day (morning and afternoon) to create intensity and urgency. Regarding contents, Week 1 completes a full introduction to the basics of MPI (Pacheco (1997)), by carefully introducing the philosophical idea of point-to-point communications between specific pairs of processes explicitly programmed in the code and then moving on to collective communications involving all processes. Week 2 tackles a larger programming problem, namely on the power method for the calculation of the dominant eigenvalue of a large matrix. This algorithm can be programmed with short, self-written code, but demonstrates both fundamental purposes of parallel computing: (i) to solve problems faster by pooling processors and (ii) to solve larger problems by pooling memory. The instruction and labs show and walk students through the idea of creating code in Matlab first to test the method in serial and using provided black-box code (for the power method), then creating serial C using by translating the Matlab, before finally parallelizing the carefully tested serial C code. By the end of Wednesday, the students set up their first large-scale simulation using their code, some runs of which will take until Friday to complete. On Friday afternoon, we introduce LaTeX and show how to create tables and include plots, with the assignment over the weekend to produce a full report on all homeworks' performance studies. This is the dry run of how to write a report as a team, including all standard manuscript components such as authors from different institutions (the participants), abstract, Section 1 Introduction, sections on the power method, its implementation and parallelization, and a summary of the numerical results, finally Conclusions, Acknowledgments (to the REU Site and HPCF for their funding), and References.

Along the way in Weeks 1 and 2, the students learn about ANSI-C style coding (Kernighan and Ritchie (1988)), the make utility, compilers, and have an introduction to scientific computing using Matlab/Octave and to statistical computing using R. To enable the fast pace of the first two weeks, we make the computer accounts available already before the participants' arrival, along with documentation about Linux in general, usage cluster for the parallel cluster specifically, and short, focused background reading on the power of parallel computing using this cluster (HPCF-2014-6; this tech. rep. is also handed out in class and used throughout the course also to teach scientific reading and writing).

The mornings of all eight Thursdays are dedicated to a GRE prep course, with pre- and post-tests in Weeks 1 and 8, respectively, and two training sessions each on reading, writing, and mathematics, respectively. During the afternoons of Thursdays of Weeks 1 and 2 starts the research component of the program. Scientists from various areas outside of mathematics and statistics spend an afternoon and evening in the program, formally presenting their problem in a 30-minute talk including questions (Figure 1) and discussion and informally being available in the break (Figure 2) and over dinner. The teams have ample opportunity to make contact with these clients and clarify their problems. Additionally, the dinners with these scientists from industry, government agencies, and other departments in attendance kicks off the professional development program by having them talk about their careers and key choices they made. This gives the students a well-rounded impression of the clients as people (useful for future interaction with them!) and of possible career choices in general.

Figure 1: Client presenting potential project.

Figure 2: Client fielding questions during break.

Phase II – Weeks 3 to 7: By the end of Week 3, each team presents their workplan for the chosen project first to all participants and mentors and then delivers it to the client. This is based on many intensive meetings of the team with its graduate assistant and faculty mentor, possibly some more lectures and labs, and possibly more communications with the client to clarify objectives and limitations. Phase II of the program over the next several weeks is fundamentally unstructured team work on the projects. It typically includes one or more updates to the client (Figures 3 and 4). These meetings always include the graduate assistant and faculty mentor and teach the students how to prepare for such a meeting by first updating just the local assistant and mentor. These are basics of project and team management, whose skills are also very useful in the context of academic research.

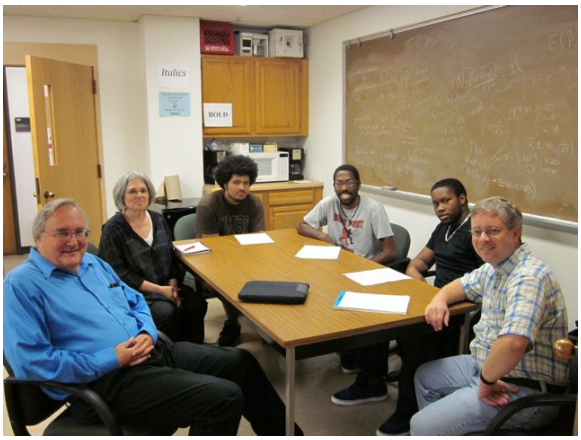

Figure 3: 2011 Team 4 meeting with client.

Figure 4: 2011 Team 2 conference call with client.

In order to continue to have some structure in the program and for the overall cohort to continue to gel, we continue the professional development program on a range of subjects that either prepare for Phase III or that deal with vital aspects of professional development, see Higham (1998) for ideas how to present these. Examples of the former include a discussion of structure of typical journal papers, the submission/review process of publications, and LaTeX as tool for making effective visual aids for oral presentations and for poster presentations. Examples of the latter include issues of professional integrity in the sciences, proper use of references, and intellectual property and copyright issues. During a few of the morning meetings at 9:00 a.m. during this phase, we welcome VIP visitors to our program, such as University President, Provost, the College Dean, and others. At each visit, along with introductions, we ask one student from each team in turn to explain their team's project to the visitor. This clearly hones

the skill of an 'elevator pitch', but it also prepares for the technique used to present a poster to passers-by. This is an example of how we use the complete set of highly integrated professional development activities throughout the 8 weeks to support the principal educational and research goals of the REU Site.

Also during the Thursday afternoons of Weeks 3 to 7, we have some talks on pertinent topics such as the Dean of the Graduate School on advice to apply to graduate school (Figure 5) or by the organizer of the UMBC Summer Undergraduate Research Fest (SURF) on tips for writing abstracts and presenting posters (Figure 6). This latter example gives an idea that some of the lectures and other activites are actually extremely carefully scheduled: First of all, by having the organizer of SURF, where the students will present their results in Week 8, meet us in person shows the students that a 'real' person is behind the SURF webpage. Then, this talk is timed to be before the abstract submission deadline for SURF, so letting the organizer show us the template and her opinion on an abstract is useful and gives the chance to clarify fine points (how several team members submit one abstract, how oral presenters are competitively selected, etc.). A lecture also by us will eventually discuss how to write an abstract (first in 100, then in 300 words), followed by a supervised lab to actually write it in submission-ready form.

Figure 5: Grad. Dean on Grad. School application. Figure 6: Organizer of SURF on prepating posters.

Phase III – Weeks 4 to 8: The purpose of Phase III is to document and present the results of the project work in all typical ways of scientific presentation to give the students the full range of experience and to make a final hand over of the results to the clients. The goal is to have a complete technical report by the end of Week 8. To make this possible, we teach the students how to start early by creating the document along the way: The starting point is the work plan at the end of Week 3, which includes the students' own formulation of how they understood the problem; the client gives feedback on this, and a correct section on the background of the problem is available in Week 4; the abstract that is due for SURF around this time provides an excellent vehicle to summarize this and give a short pitch of the proposed solution technique. This is extended by a section on the method used to solve it, while it is proposed in Week 5. In Weeks 6 and 7, we guide the students how to manage and present potentially large amounts of data generated; clearly, this is a give-and-take with modification of the method, additional studies, and new proposed conclusions --- *the students experience real research, since also the assistants, mentors, and clients all do not know the result!*

We feel that it is absolutely vital to also show students how to bring closure to a project in a limited time frame. We accomplish this during Week 7 by using a multi-step process from creating slides for an oral presentation that include only bulleted lists, to the poster that forces

the design of one appropriate summary table or plot together with few well-chosen full sentences, to a project webpage for each team that contains the material of the poster but adds more details to the paragraphs. In Week 8, this carefully edited webpage is an excellent starting point for completing the most important sections of the team's tech. report, the Introduction and the Conclusions. With this integrated and guided multi-step process, we are able to have an extended abstract, a webpage, slides, a poster, and a tech. report completed by the end of 8 weeks and exposed the students to the craft involved in documenting research. Week 8 also includes other wrap-up activities such as delivery of any promised products to the client (e.g., code or results) and clean-up and turn-over of the team's directory on maya, where all code, poster, report, and other material is collected, to the mentor.

Figure 7: 2011 Team 2 with faculty mentors.

Figure 8: 2011 Team 4 with client and faculty mentor.

The highlight of Phase III is certainly the public presentation of their work at the UMBC Summer Undergraduate Research Fest on Wednesday morning of Week 8. In recent years, there have been over 100 posters presented at this university-wide event, organized by the College of Natural and Mathematical Sciences. Our program is recognized as a regular participant. The students enjoy the event greatly and all clients who are able attend the event (Figures 7 and 8). SURF also includes one oral session, which usually has only five presentations that are selected competitively from the abstracts submitted. In recent years, our program has typically provided two of the oral presentations.

Some of the VIP visits over the previous weeks are also specifically meant to showcase role models, such as members of underrepresented minority groups in campus leadership positions. This is complemented by a discussion of history of computing, which notably includes interesting aspects on the role of women in the mathematical sciences, see Grier (2006) and Zitarelli (2006), which we make available and discuss. We also start a conversation about publication opportunities for the project work and discuss with students the option for presenting their work orally or as poster at other conferences. Some participants have presented their work at their home institution in departmental seminars or institution-wide sessions on undergraduate research. Others have given oral or poster presentations at major national conferences, such as the Joint Mathematics Meeting or the SIAM Conference on CS&E. Several teams have successfully converted or extended their tech. reports into publications, for instance in the web journal *SIAM Undergraduate Research Online*. Whenever we become aware of such success stories, we take the opportunity to post them as news items on our REU Site webpage.