

January 31, 2006

OUTLINE: SOE DEPARTMENT/PROGRAM STATEMENTS FOR REVISED ACADEMIC PLAN

Ira Pohl
Computer Science Department

- 1. Maintaining and Building Excellence**
- 2. Sustainability within Available Resources**

Computer Science Program Statement

The quality of UC Santa Cruz's Computer Science Department is reflected by the accomplishments of its faculty. UC Santa Cruz's Computer Science Department has been in existence for over thirty years and offers four degrees: B.A., B.S., M.S. and Ph.D. in Computer Science, with a combined BS/MS degree currently under development. The program includes nineteen full-time ladder rank faculty with over 130 students enrolled in the graduate program. To date, the Department has awarded more than 220 MS and 80 Ph.D. degrees. University-industry interaction is enhanced through the employment of computer professionals as visiting faculty and through arrangements for students to gain practical research experience by working as interns in nearby industrial research laboratories.

Computer Science Department Faculty

The Department is highly regarded on campus and at the national level for its excellent faculty, extramural funding, and high quality of teaching. Of the current nineteen faculty members, ten are full professors, three are associate professors and six (one of whom was hired this year) are assistant professors. The technical strength and the impact of faculty research is demonstrated by their appointments to the editorial boards of several ACM and IEEE journals, a Sloan Foundation Fellowship, two ACM Fellows, an IEEE Fellow, the last four junior faculty received NSF CAREER awards, and participation in numerous technical program committees and NSF panels. Computer Science faculty members have funded research projects and publish regularly in leading technical journals. The Computer Science faculty received \$2,624,000 in 2004–5 and \$11,589,000 during the past five years in extramural contract and grant funds from federal agencies and industrial sponsors.

Computer Science Research

The Computer Science faculty conducts research in following primary areas: Computer Graphics and Scientific Visualization, Computer Systems, Machine Learning, Databases and Software Engineering. The following paragraphs provide overviews of these areas.

The Department of Computer Science is highly regarded for its strength in **Computer Graphics and Scientific Visualization**. Three faculty members and more than 25 graduate students are actively engaged in this area of research. Research in scientific visualization is funded by several agencies including NSF, DARPA, DOD, and ONR. The faculty works in close cooperation with investigators from the different scientific and engineering disciplines. Particular software interests currently include algorithms for volume rendering and isosurface generation from volumetric data, user-interface issues, uncertainty visualization and

parallel processing as applied to visualization. In the field of computer graphics, the faculty is actively engaged in physical simulation. Physical techniques are useful for simulating rigid articulated bodies representing robots, humans, and other animals, and flexible bodies such as cloth, skin, and other elastic materials. The group is interested in exploring faster algorithms for physical simulation, in the hope of making it interactive. Constraint and control methods are also being investigated. The group is utilizing a variety of sensors: GPS (Global Positioning System), orientation trackers, inertial sensors, cameras, camcorders, stereo cameras, and LiDARS to create a four-dimensional space-time visualization of geospatial data and environment-related intelligence.

The Department enjoys great strength in the area of **Systems Research** with seven faculty members and 20 graduate students working in storage systems, distributed computation, programming languages, and database systems. In storage systems, researchers are investigating key challenges in storage including storage scalability to petabytes and beyond, new storage technologies, and long-term archiving. Through the UCSC Storage Systems Research Center, storage researchers at UCSC receive significant funding from the NSF, the National Laboratories, and industry sponsors. In distributed systems, researchers are investigating how the topology of the network and the knowledge available to individual processors affects the computational efficiency of a system of distributed processors and what functions can be computed in a network where the processors are anonymous. They also study the use of formal logics for reasoning about distributed computations. The programming languages group is focusing on three sub-areas: object-oriented programming, parallel programming and logic programming. They are studying the object-oriented programming methodology using the C++ language and are interested in the extension of language with concurrency and in providing automatic storage reclamation. A current research project at UC Santa Cruz involves the investigation of techniques for analyzing parallel programs and their associated execution traces. The analysis can be used to aid in the parallelization of a program, in understanding the program, and in debugging the program. An important part of the project is the development of a graphical browser for viewing the results of the analysis. The group is also engaged in the implementation of a powerful logic-programming compiler. The computer systems group is researching a range of systems-design issues focusing on practical implementation based on sound theoretical foundations. Examples of interest include real-time systems, computers as system components (embedded systems), special purpose processors, and digital networks and their associated components. Other research in this field includes work in performance prediction, evaluation, and optimization, since these are major tools for use in systems design and modification.

The **Machine Learning** group at UC Santa Cruz is primarily interested in the theoretical aspects of machine learning; in particular it has been at the forefront of research in Computational Learning Theory and has hosted many of the major conferences in this field. Over the past eight years they have been developing a new family of on-line learning algorithms with qualitatively different behavior than the previously known gradient descent techniques. Work is in progress on extending this family of algorithms and quantifying the performance differences between these algorithms and existing algorithms in various settings. One of the most successful practical techniques for machine learning is boosting. The most successful boosting algorithm was derived from the family of on-line learning algorithms developed at UC Santa Cruz. Furthermore, many of the techniques used to analyze this and other learning methods were developed or introduced at UC Santa Cruz. This group has produced some excellent graduates who have gone on to successful careers in industry and academia, most notably Yoav Freund and Nick Littlestone, whose on-going contributions to boosting and on-line learning, respectively, are some of the most important in the field. After finishing Ph.D., Yoav Freund held a research position at AT&T Bell Laboratories and is now at UC San Diego. Nick Littlestone held a postdoctoral position at Harvard University.

Software Engineering research at UC Santa Cruz is performed by three faculty members, whose research is primarily supported by the NSF. Software engineering research focuses on developing tools and techniques for improving the quality and evolution of large software systems. Research on model checking of software

components and their composition promises advances in analysis of systems at an architectural level. Object-oriented design and design patterns research improves our ability to represent and reason about software during the design phase. Projects focused on pair programming in the classroom, literate programming technology, and software configuration management repositories aim for improvements in the activity and control of software coding. Research on object-oriented methodology using contemporary languages such as C++, Java, and C# yields improved understanding of how to represent complex problems in software. Static analysis and verification research leads to improved understanding of software instabilities within large evolving systems, as well as improved quality in new and existing software.

Database Systems research is led by three faculty members and includes five Ph.D. students. There are four primary research areas, namely, schema mappings and data exchange, data provenance, data reduction and approximation techniques, and self-organizing systems. Research on schema mappings and data exchange focuses on the investigation of foundational techniques for the management of schema mappings that specify the data exchange or relationships in a network of inter-related databases, as well as answering queries posed in such a network. In the context of data provenance, the focus is on the investigation of foundational techniques as well as on the development of effective tools for tracing the provenance and flow of data in such a network. Research on data reduction and approximation explores techniques for the effective summarization of large data stores and the computation of approximate answers for complex queries. Finally, the project on self-organizing systems looks into the problem of building autonomous database systems that can adapt automatically to the characteristics of their operating environment. These research projects are supported primarily by the NSF and gifts from IBM and Microsoft, and involve collaborations with researchers in both industrial labs and other universities, in particular, IBM Almaden Research Center, Intel Research, University of Toronto, University of Tel Aviv, and I.N.R.I.A. in France.

3. Future Opportunities for Investment in new Endeavors

4. Synergistic Graduate Programs

New Programs and Interdisciplinary Research

UC Santa Cruz's Computer Science Program also has a strong interdisciplinary focus, conducting research with researchers in others areas and helping create new interdisciplinary programs. Until the recent founding of the new Biomolecular Engineering (BME) Department, UC Santa Cruz's Bioinformatics Program was housed within the Computer Science Department, and a number of current Computer Science graduate students continue to pursue degrees in this area. The Department is currently housing the UC Santa Cruz Technology and Information Management (TIM) program. Finally, several researchers in the Department were recently provisionally awarded \$3,750,000 from Los Alamos National Laboratory to create the new UCSC/Los Alamos Institute for Scalable Scientific Data Management.

The **Institute for Scalable Scientific Data Management** (ISSDM) is an education and research collaboration with Los Alamos National Laboratory (LANL). The ISSDM focuses on Scientific Data Management advanced research and development topics in the areas of simulation and real time/experimental data collection, storage, analysis, and organization management. By working collaboratively with industry partners, LANL, and UCSC, the Institute will help solve Simulation Data Management problems at unprecedented scale and of national importance while also supporting CCN division, DOE/NNSA Advanced Simulation and Computing (ASC) program, and LANL Institutional Computing program goals.

The Computer Science Department emphasizes the placement of its degree recipients and actively assists them in obtaining rewarding positions. Graduates have gone on to a variety of positions in academia and industry. A number of computer science graduates have pursued teaching careers, securing positions at

institutions such as Rice University, Johns Hopkins University, the University of Pittsburgh, UC Berkeley, and a recent position at UC San Diego. Placements in industry have included positions at bellwethers such as Apple Computer, Bell Labs, IBM Almaden Research Center, Micron Technology, National Semiconductor, Oracle, Raytheon Corporation, Sun Microsystems, Sarnoff Corporation, SGI, Veritas, Xerox, Yahoo, and several startup companies. Additionally, some students have accepted employment at government research facilities, including Los Alamos National Laboratory, the Naval Research Laboratory and nearby NASA Ames Research Center. The strong placement record the Computer Science Department has compiled is not only a reflection of the strength of its programs, but also the quality of its students.

Computer Game Engineering

We are recruiting in computer gaming in order to create a new sub-major. This is important in two regards: computer gaming is a fast growing research area that integrates existing strengths in the department; computer gaming is very attractive as a recruitment tool for very high quality undergraduates. Computer gaming already exists as a pathway in the ordinary CS major. By creating a new named degree program tentatively titled Computer Game Engineering we expect to reverse the recent decline in number of CS majors and increase the quality of entering freshmen. UCSC will have the first such degree in the UC system that emphasizes rigorous technical computer science.

Resource issues for this new degree are adequate for accepting 25 majors or slightly more each year. The courses within CS for this degree already exist or can be managed when the new CS gaming position is filled in 2006. Programs that have been contacted by us, especially digital media, economics, music and mathematics, all welcome this initiative and readily expect to accommodate the first cohort.

Plan for additional faculty FTE, with specific areas of concentration identified.

Computer Science Hiring

We have grouped our hiring priorities into four tiers: principal, priority, lower priority, and other.

Principal:

1-2) Positions supporting the Computer Gaming initiative. This initiative is attracting great interest from undergraduates and development is progressing forward rapidly.

Recruiting position 1 in AY2005-2006. Recruitment of position 2 is AY2007-08.

3) Position supporting Databases. The Database Systems group is becoming of increasing importance and is essential for the planned SOE excellence in Information Technologies. An additional hire will enable us to have critical mass in this area.

Recruitment of position 3 in AY2008-09.

4) Position supporting Software Engineering/Programming Languages. Software Engineering has been a planned growth area in the SoE for many years and with the recent departure of Raymie Stata to industry, we need another faculty member to sustain our critical mass.

Recruitment of position 4 in AY2008-09.

Priority:

5) Position supporting Machine Learning/Data Mining. As with Databases, the machine learning area is important for SOE's excellence in information technologies. Furthermore, many disparate

areas within the SOE are finding that machine learning techniques are useful and often essential tools for their own research.

Recruitment of position 5 in AY2009-10.

6) Position supporting Operating Systems. We have a very important storage systems group of three faculty. It is very desirable to build out from this strength by hiring new faculty with related interest.

Recruitment of position 6 in AY2009-2010.

7) A third position supporting Computer Gaming/Entertainment. The third hire in this area would enable it to achieve critical mass.

Recruitment of position 7 in AY2010-2011.

Lower priority:

8) Position supporting Computer Security. Security has long been an area of interest to the computer science department, but we feel that expanding to this area is a lower priority than the above positions. However, it would be delighted to hire an outstanding candidate in this area.

Recruitment of position 8 in 2010-2011.

9) Position supporting Computer graphics. We would like to build on our excellence in computer graphics. This is a lower priority as we would like to see if the computer gaming hires provide adequate support in this area before running a focused recruitment. This would be a replacement for Jane Wilhelms.

Recruitment of position 9 in 2010-2011.

Other

10) Positions supporting Algorithms/ML/AI/Systems/SE/other. These are areas in which an additional faculty would be welcome. It may be that the ML and gaming initiative positions will provide adequate support in the fundamental area of algorithms. If current trends continue, additional faculty in the Machine Learning area would be desirable. Similarly for systems and software engineering, the previous hiring's can also provide an opportunity to continue expansion in an as yet undetermined direction. It might also be the case that this FTE could be split with another program, such as TIM within SOE, or Digital Media in the Arts. This would allow two shared appointments.

Recruitment of position 10 TBD.

Plan for enrollment FTE—both undergraduate and graduate students.

The Computer Science department currently has 270 undergraduates (declared majors) and 140 graduate students.

We expect to bring on board approximately 25 or more new CS gaming students a year over the next 4 years. The anticipated minimum additional enrollments of CS gaming students: 2006 – 25; 2007 - 50; 2008 – 75; 2009 - 100 (steady state). We also expect an upturn in the major because the CS job market has recovered from the dot-net bubble bursting.

We expect to increase our service offering in several ways over the next two years. We have a new lower division computer gaming course CMPS80k that is expected to attract 200 students (preliminary enrollments

statics support or optimism). Our general education programming courses are gaining in popularity and are required by the business economics major and some science majors. We expect to have 300 students in our yearly offerings for these courses. CMPS 10, a feeder and general education course, remains very popular and has two offerings with nearly 300 students per year. We expect the graduate program to grow in proportion to faculty growth—namely approximately 7 grad students per faculty.

Plan for extramural research support in specific areas.

GAANN, Sloan, Career Awards, Science watch measures, extramural grant and gift support.

CS Department faculty were recently awarded \$3,750,000 from Los Alamos National Laboratory to create the new UCSC/Los Alamos Institute for Scalable Scientific Data Management.

Gift support has continued to be very strong.

GAANN has supported up to 7 graduate PhD students.

1. Interdepartmental and Interdivisional collaborations
 - a. Computer Gaming involves interdivisional collaboration with the arts, especially the digital media program which itself is interdisciplinary.
 - b. The SSRC has members from other engineering disciplines.
 - c. The machine learning group collaborates with AMS and Economics. Machine learning algorithms are foundational in several disciplines.
 - d. The graphics and Visualization group have collaborations with CE.
 - e. The software engineering group has collaborations with CE and TIM.
2. Diversity
 - a. GAANN award for CS graduate research, Department of Education, \$250,000 annually for three years. Among the 7 recipients are 4 women and 2 Latino's.
 - b. There is a Cota Robles fellow who is a Latina.

Estimation of overall support is difficult to extrapolate. Based on current department averages, we would expect by 2011 to be generating at least 3.6 million in extramural support. If major initiatives and inflation are added in, this ups the projection to between 4.5 – 5 million dollars annually.

Recent Computer Science Department Highlights

- Computer Science hired James Davis in 2004, who is in graphics systems. This greatly aids our visibility in graphics, visualization and potentially animation and gaming. Computer Science has hired Dimitris Achiloptas in theory. He has been awarded an NSF Career award for \$400,000.
- Cormac Flanagan received the Sloan Fellowship, the only one awarded to UCSC in 2005.
- The Malvalli Chair was established and Darrell Long was its first recipient. He was made an IEEE fellow in 2005.
- The database lab ran a weekly seminar and has achieved critical mass and national visibility. Wang-Chiew Tan received a Career award from NSF. Alkis Polyzotis has received a Career award from NSF and has an IBM fellowship. Phokion Kolaitis, on leave to IBM Research, has been made an ACM Fellow in 2005.

- We have continued to expand our corporate support. We have initiatives with Microsoft Corporation to include them in support of our CS research. Microsoft has included us as one its research university partners. It has involvement or is funding Abadi, Long, Pohl, Flanagan, and Polyzotis. This funding spans several research areas of importance to our program: security, programming languages and software engineering, storage systems, and database systems. The SSRC is especially well funded by a large number of companies, including IBM, Microsoft and HP.
- Computer Science has expanded its curriculum offerings in SE and Database at the advanced undergraduate and graduate levels. It has modified its comp requirement and has moved to a capstone model for undergraduates. CEP accepted Cmps116 , CMPS161, CMPS 181, CMPS 183 and CMPS 140 as such courses. A similar offering will be available in the computer gaming area.
- Five Computer Science Faculty appear on the Science Watch List of the 250 most cited CS researchers in 2004. This is the most of any UC campus and this is the case where the average UC CS department is approximately 34 faculty versus UCSC with 19 faculty.