

Evaluating Research of USNA CS Department Faculty

Research Committee of the Department of Computer Science

November 20, 2007

1 ABSTRACT

The research records of USNA faculty are evaluated for P&T purposes by an interdisciplinary, yard-wide committee. Evaluating research is always a difficult enterprise, but under these circumstances it is made especially difficult by differences in how research is carried out and, most importantly, communicated in different disciplines. In this document, we in the Department of Computer Science attempt to describe the unique practices of our discipline and the issues these practices raise in the evaluation of CS research records. We hope that it will serve as a resource for the Dean and members of the P&T Committee.

The principal points are:

1. In computer science, a publication in the proceedings of a reputable conference should be considered on par with a journal publication. It undergoes a similar level of peer review, and it is just as much an “archival” publication.
2. Point (1) is recognized in the field of computer science, endorsed by professional societies, and acknowledged by the P&T process at many other institutions — including all but one of the schools responding to our Department’s queries on this topic.
3. Software can constitute an important research contribution and should not be ignored in the P&T process.

2 PURPOSE

In the 1990s many CS Departments put a lot of work into educating the administrations of their institutions on research practices in the field, because research that was important and influential to the CS community was not recognized as such by the traditional metric of academia: a count of journal articles. In this context, the Computing Research Association (CRA) — an association of more than 200 North American academic departments of computer science, computer

engineering, and related fields; laboratories and centers in industry, government, and academia engaging in basic computing research; and affiliated professional societies — published the white paper “Evaluating Computer Scientists and Engineers for Promotion and Tenure”, Enclosure (1) of this memo. This article, written by three prominent researchers in the field, argues eloquently and authoritatively that scholarly achievement in computer science should not, indeed cannot, be evaluated exclusively through journal publications. In our canvassing of other academic institutions for input, the CRA paper was explicitly recommended by Associate Deans at the University of Wyoming and Ohio State, and CS Department Chairs at Villanova and Drexel University. We respectfully submit it for consideration by the Dean and the P&T Committee.

At our institution, of course, P&T decisions usually follow the recommendations of an interdisciplinary committee with rotating membership. Members of this committee are tasked with evaluating the research records of CS Department candidates. A primary concern in our department is that the yard-wide P&T committee, with its rotating membership, is not always well-informed on CS research practices and how CS research should be evaluated. In particular, the committee:

1. usually has no representative involved in CS research,
2. does not have the institutional memory to retain familiarity with CS research practices from year to year, especially since CS faculty come up for promotion infrequently, and
3. may discount Department Chair input on CS research practices and the evaluation of CS research because the Chair may be seen as an advocate for the candidate.

This memo and its attachments can serve as a resource for future Deans and P&T committees to assist in evaluating candidates from the CS Department; providing “institutional memory” and guidelines that are independent of any particular Department Chair or candidate under consideration.

3 WHAT’S DIFFERENT ABOUT CS

The CRA white paper’s introduction outlines the fundamental issues:

“The evaluation of computer science and engineering faculty for promotion and tenure has generally followed the dictate ‘publish or perish,’ where ‘publish’ has had its standard academic meaning of ‘publish in archival journals’ [Academic Careers, 94]. Relying on journal publications as the sole demonstration of scholarly achievement, especially counting such publications to determine whether they exceed a prescribed threshold, ignores significant evidence of accomplishment in computer science and engineering. For example, conference publication is preferred in the field, and computational

artifacts — software, chips, etc. — are a tangible means of conveying ideas and insight. Obligating faculty to be evaluated by this traditional standard handicaps their careers, and indirectly harms the field.”

In terms of our department, the issues can be distilled into:

1. papers published in the proceedings of reputable conferences need to be considered as on par with journal publications, and
2. software can constitute a research contribution that should be considered for P&T purposes.

The first of these is the more important, as it applies to all our faculty members, and it is the focus of this memo.

4 CS CONFERENCES

The role of conferences and conference proceedings in computer science is different than in many other disciplines. In particular, for reputable conferences in computer science: 1) submissions are full papers, not abstracts, 2) submissions undergo thorough peer-review and 3) the proceedings in which conference papers are published are “archival” — they are part of library collections and the papers therein are the cited building blocks of further results by other researchers. In short, papers published in the proceedings of reputable conferences are peer-reviewed, archival research artifacts just like journal articles, and need to be evaluated as such.

To clarify points 1-3:

1. Full-paper submissions: Full-paper submissions are, in some sense, one of the defining features of the conferences we call “reputable”. As examples, call-for-papers from conferences at which our faculty members have made presentations can be furnished if required.
2. Peer-review: Papers submitted to reputable conferences are subject to a level of review that is comparable to, and which sometimes exceeds, that of journals. Examples:
 - Papers submitted to the ACM Sponsored ISSAC2007 were reviewed by an average of 3.5 external reviewers in addition to reviews by program committee members and discussions during the program committee deliberations.
 - The review process for ISWC 2006 (Int. Semantic Web Conference): The review process included three distinct phases. First, all papers were evaluated by three members of the Research Track Program Committee. Then, each paper and associated reviews provided the basis for the meta-review phase, led by an experienced member of

the Program Committee who had not participated in the first phase. This strategy produced a joint recommendation from reviewers and the meta-reviewer to the Research Track Program Co-chairs, who, in a final review phase, analyzed each recommendation in detail, in some cases commissioning additional reviews and initiating further discussions. The Program Chairs then made a definitive decision regarding each paper.

- The review process for AAAI 2007: Every paper received at least three reviews. There were over 500 program committee members. The reviews and often the papers themselves were read by one of 45 senior program committee members, and a discussion among reviewers was initiated. Author feedback was then taken into account for the final discussion, and when necessary, additional reviews were collected. Discussions were often very detailed and, in most cases, led to resolution of issues brought up by the reviewers.
 - The review process for the IEEE ICSM 2006 (International Conference on Software Maintenance) had each submission reviewed by at least three external reviewers plus the discussion and review during program committee deliberations.
3. Archival: The point of “archival” research publications is that they provide the building blocks upon which future work is based. Thus, to demonstrate that conference proceedings publications are as “archival” we chose indisputably important journal and conference proceedings papers from various areas of CS¹ and compared the number of journal citations to conference proceedings citations in their bibliographies. There were variations across the disciplines, but on average there were more conference proceedings publications cited than journal articles by a ratio of 1.5 to 1. Enclosure (3) presents the full results of our survey, summarized below:

Subdiscipline	#conf.	#jour.	ratio
Artificial Intelligence	99	50	1.98:1
Machine Learning	28	44	0.64:1
Databases	174	63	2.76:1
Graphics	162	161	1.01:1
Semantic Web	71	25	2.84:1
Software Engineering	76	44	1.73:1
Symbolic Computation	23	49	0.47:1

Clearly conference proceedings publications in CS are every bit as “archival” as journal publications.

¹Some papers were Distinguished Paper winners at conferences, two were most downloaded papers from a Journal’s website, one won an award as the Most Influential Paper of the Last 10 Years, etc.

5 HOW THINGS WORK AT OTHER INSTITUTIONS

“Research conferences are often the most desirable venues for presenting our research results.” – David A. Patterson, 2004, as President of the Association for Computing Machinery

In preparing this memo, members of the Department contacted current and past department chairs, deans and P&T committee members at other colleges and universities for information about the issue of conference proceedings in computer science and its handling at their institutions for promotion and tenure decisions. We received replies from top-tier institutions like Stanford, MIT, Maryland and Ohio State, as well as other strong schools like Drexel, Notre Dame and Villanova. Their responses² show that other institutions, including prestigious institutions with masters and PhD programs, treat strong conference proceedings publications as comparable to journal publications. Clearly, institutions such as USNA should hold conference papers in no less esteem than premier research institutions. We quote sections of a few of the replies we received to highlight this fact:

“In CS, we consider a paper in a top-rated conference (e.g., ISCA for computer architecture or FOCS for theory) to be more valuable than a journal publication. This is because (1) the top conferences are more selective than journals [...], and (2) top conference publications have greater impact. More people read these conference papers than read journal papers - even 10 or 20 years later.” – Bill Dally, Chair of the Department of Computer Science, Stanford.

“I know it seems strange to people outside of our discipline, but conference publications - in the right highly selective conferences - are given far more weight in tenure deliberations within CS than journal publications!” – Larry Davis, Chair of the Department of Computer Science, University of Maryland.

“We (meaning both the EECS department and Engineering Council – all the engineering department heads plus the dean, a group that oversees promotion cases) routinely judge promotion cases based on the publishing criteria of the field. In most areas of computer science, this means that elite conference publications are much more important and expected than journal publications.” – Eric Grimson Bernard Gordon Professor of Medical Engineering, EECS, MIT Head, EECS, MIT

“In most areas of computer science certain conferences and a few journals are the most influential, competitive and prestigious publication venues” – Susan Egger, Professor and former P&T Committee Member at the University of Washington

²These correspondences are available if needed.

“In some of the major theory conferences, for example, FOCS and STOC, many of these papers never get published as journal papers. The conferences are actually more prestigious than the journals. Once published in FOCS or STOC, the research is sometimes never published in another venue” – Raymond Greenlaw, Dean of the School of Computing, Atlantic Armstrong University

“All conferences are not created equal, but it is well known that in our discipline the strong conferences have both review processes and impact comparable to good journals.” Stu Zweben, Associate Dean, Ohio State

“major conference publications are counted as journal publications in my report [for P&T nominations]” Jeremy Johnson, Chair of the Department of Computer Science, Drexel University

6 DESPITE USUAL CS PRACTICES, WHY NOT JUST PUBLISH IN JOURNALS?

The above is a statement of *how* things are in computer science without addressing *why*. It is natural to ask: Why not just publish in journals?

1. First and foremost, in many areas of CS the most prestigious venue for publication is the proceedings of a conference, as several of the quotes from the previous section support. P&T requirements should not dictate publication venues to faculty — especially to the detriment of their stature in the research community.
2. Results published in a conference proceedings are also presented at the conference, which can make the work much more visible to the community. In this context, it is important to understand two things:
 - (a) Work published in proceedings usually can't be republished in a journal, as the community considers that the work is already “published”. To turn around and publish the result in a journal, either the paper has to add a substantial amount of new material to the conference paper, or the paper would have to be a retrospective survey of a result developed over several conference papers. The former can actually be disadvantageous: the new result added onto the conference paper work can be obscured by the previously published material it appears with.
 - (b) Work published in a journal cannot be presented at a high-level conference. The only presentations at a conference are the invited talks — big picture talks by eminent scientists — and presentations of papers from the proceedings. If a paper has been published in a journal, it cannot be republished in the proceedings, so you can't talk about it to the community.

3. The turn-around time for conferences is much shorter than for most journals, which is very important in our rapidly changing field. From submission to publication is usually about six months. It might be two years or more for a journal. This does not mean the quality of review is less. However, it does mean that a paper that the referees and/or program committee deems to require substantial revisions is rejected out of hand, even if the results are fundamentally sound, whereas with a journal there are often several rounds of revisions and resubmissions.

7 SOFTWARE ARTIFACTS

Certainly journal or conference proceedings papers are not the only artifacts of research. In computer science particularly, software can constitute an important research contribution. This can be seen very clearly in the choice of “Turing Award” winners by the ACM. The ACM Turing Award recognizes individuals who have made lasting and major contributions to computer science, and is probably the most prestigious award in the field. Consider the citations for awards given this millennium:

- 2005 Naur, Peter - Design of the ALGOL programming language.
- 2004 Cerf, Vinton G. and Kahn, Robert E. - Development of TCP/IP the stack of software layers underpinning the internet.
- 2003 Kay, Alan - Design of the Smalltalk programming language.
- 2001 Dahl, Ole-Johan and Nygaard, Kristen - Design of the Simula programming language.

It is tempting to equate software created by a researcher in computer science to an experiment from another discipline. However, there are fundamental differences that are very important for the evaluation of research.

1. An experiment is performed or a system is built, and one would expect to write a paper about it. However, when an experiment is completed, there’s no further work to do. A computer program is always under development, bugs must be fixed, new systems or upgrades require the program to be ported, many incremental improvements are made, none of which will ever make a paper.
2. The culture of CS community, as seen in the ACM/IEEE “Software Engineering Code of Ethics”, is to make research software freely available to other researchers. When a researcher’s software is used in someone else’s work, he may get a citation, he may get listed in the paper’s acknowledgments, there may simply be a footnote with a URL, or perhaps nothing beyond the name of the program appearing in the text. In some fields, if a sample collected or created by a researcher as part of an experiment

is used for some other work, he may be listed as a co-author. This is definitely not part of the culture in computer science.

Certainly not all software constitutes a research contribution. But, as discussed in the CRA whitepaper, software can represent an important research contribution and that has to be considered in evaluating a computer scientists research.

8 RECOMMENDATIONS FOR EVALUATION

The P&T committee has the unenviable task of evaluating the research records of candidates from many academic disciplines. In evaluating candidates in computer science, the committee has to make decisions about weighting the importance of each candidate's conference proceedings publications. As Dean Zweben from Ohio State noted, "all conferences are not created equal", although the same can be said of journals. We wish that there was a simple algorithm for rating the strength of a given conference, but there is not. However, we do comment on a few issues that should be considered.

First and foremost, strong conferences require thorough peer review of full-paper submissions. Information on the refereeing process is usually given in the forward to the proceedings, which candidates can make available to the P&T Committee.

Second, there are certain conferences that are recognized as the top-conferences in their areas. Some examples are mentioned explicitly in the responses we received from chairs/deans of other institutions. The P&T Committee could get advice on this from acknowledged subfield experts. This could be addressed in the chair's brief to the committee or in external letters in candidate packages. As was repeated by most of our outside sources: publications in the proceedings of these top conferences should, without question, weigh as much or more than journal publications.

Third, papers from conferences with full-paper submissions and thorough peer-review processes that are not the "top" conferences in their fields are harder to evaluate. Acceptance rates and sponsorship by professional organizations like the ACM or IEEE are indications of a publication's quality, but they don't tell the whole story. AAAI is not affiliated with either ACM or IEEE, but is the top conference in artificial intelligence. The ACM sponsored ISSAC is the top conference in symbolic computing but, coming from a small community, it has a fairly high acceptance rate. In the end, input from subject area experts is probably the only way to determine the relative importance of a particular conference. What is crucial to keep in mind, however, is that the conferences under discussion all require full-paper submissions and a peer review process as thorough as any journal.

Evaluation of the scientific merit of software artifacts is an even more difficult proposition. What previous sections and the CRA whitepaper have hopefully made clear, however, is that software artifacts *can* represent important research contributions — they cannot be dismissed out of hand. It should probably be

incumbent on the candidate to argue that a particular piece of software is an important contribution. We simply mention citation or other acknowledgment in research papers as one demonstration of scientific importance. Others might include tutorials at conferences or invited talks related to the software.

9 CONCLUSION

Evaluation of research only makes sense within the context of the practices of the discipline in which the research takes place. Thus, evaluating the research record of a P&T candidate in computer science requires an understanding of the practices of our field. We hope that we have provided the Dean and the P&T Committee with an overview of the special role of conferences and software within our discipline in a form that will help guide sound, well-informed decisions on future candidates from our department. We have tried hard to include statements from high-visibility sources to provide what is truly the discipline's perspective, not just the Department's perspective on these issues. We acknowledge the difficulty of the job the P&T Committee has to do, and hope what we have produced will be a valuable resource.