What happens to a Computer Science research after it is published? Tracking CS research lines*

Jacques Wainer Eduardo Valle

July 27, 2012

Abstract

How many Computer Science (CS) papers are extended after they are first published? That is the central question of this work, based on a statistical survey of two populations of papers, one from journals and one from conferences of the area. We sample 100 papers from each population, and use self-citations to identify potential and actual continuations. We are interested in asking how many papers do indeed continue, how and when does the continuation take place, and what are the differences (if any) between the journal and conference populations. Despite the implicit assumption of a research line behind each paper, manifest in the ubiquitous "future research" notes that close most papers, we find out that more than 70% of the papers are never continued.

1 Introduction

A sketch of future research directions has become a customary closing for research papers in Computer Sciences. That practice reveals an implicit, but pervasive assumption of scientific production in our field: that behind each article there is a line of research, which persists after publication is accomplished. We suspect, however, that many authors never fulfill that hint at continuation, and that many papers are the terminus of their research lines. How many do effectively continue? Does it make a difference whether the paper was published on a journal or conference? Does the chance of continuation correlates with the paper reception by the community?

Those are some of the questions we are interested in answering. Using a statistical survey as a methodological tool, we sampled independently journal and conference papers, and used self-citations to those papers to explore how

*This version is a preprint of an article accepted for publication in the Journal of the American Society for Information Science and Technology (JASIST). The final version, edited by the publisher is yet to appear.

they were extended by their authors. The underlying assumption is that, in a single research line, follow-up papers will cite the previously published ones, making self-citations useful to track the continuations.

By contrasting the journal and the conference samples, we also intend to contribute to the ongoing debate about the roles of those media on the field of Computer Sciences. Is research more likely to continue when it is published on conferences, and thus do conferences play the same role as in other disciplines, as places to report ongoing efforts and receive feedback from the community? Or is publication in conferences actually *less* likely to continue, suggesting that our conferences are places to publish less elaborated, more speculative (and thus riskier) work?

We are interested in not only quantifying how much of the work in CS continues after publication, but also in how does that continuation proceed. Specifically, we want to know:

- how many research lines continue after that publication?
- how do the continuations split between journals and conferences?
- how long does it take for the next paper in a research line to be published?
- who pursues the continuation — roughly the same group, or just one author? Are the senior or the junior researchers mostly leading the continuations?
- is there a correlation between the impact — measured by the citation count — of the original paper and the fact the research line continues?
- is there a correlation between the quality or selectivity of the original medium and and the fact the research line continues?

For each of those questions, we are interested in the differences that might exists between originals published on conferences or on journals; and what those differences tell us about those two media.

We have sampled 100 journal and 100 conference papers published by ACM in 2003, which we have called the journal and conference **originals**. We then collected, in November of 2011, the list of papers that, according to ACM Digital Library (ACM DL), cited the originals. If any of the original authors were co-authors of the citing papers, we have considered that the citing paper was a potential continuation or a **follow-up** of the original research line. We have then analyzed all the follow-ups to verify if indeed the research published was truly a continuation of the original. In that case we say that the follow-up is an **extension** of the original papers.

2 Related research

2.1 Conferences and Journals

In the discipline of Computer Sciences, the role of conferences and journals as publication media has been a contentious subject (Fortnow, 2009; Grudin, 2011; Freyne et al., 2010; Halpern and Parkes, 2011; Wing and Guzdial, 2009; Birman and Schneider, 2009). In most scientific fields, conferences aim at promulgating works in their earlier development stage. Therefore, the review requirements are lower, and the proceedings, when they are produced at all, are not intended for archival and reference use. That is conspicuous in the much lower citation count conference proceedings receive, when compared to journal articles: proceedings attract less than 5% of all citations, a fraction that has been steadily decreasing (Lisée et al., 2008).

However, in Engineering, and especially in Computer Sciences, the situation is not obvious. The traditional evolutionary model from early drafts on conference proceedings to mature work on journal articles is no longer clearly observable. In stark contrast to other disciplines, proceedings get a large — and growing — fraction of all citations (Lisée et al., 2008).

The reasons for that relatively privileged status of conferences in Computer Sciences are not completely understood. Some authors have pointed to the accrued importance of timeliness in a rapid evolving field such as Computer Sciences, making conferences preferable (Montesi and Lago, 2008). However, the emphasis of the field on conferences has come under growing criticism, on the grounds that it leads to hasty peer-review, deadline-driven research, and deprivation from conferences of their usual role of fora for debate and exchange of ideas (Grudin, 2011; Wing and Guzdial, 2009).

It is clear that the classical evolutionary model fails to account for the publication conventions of Computer Sciences. What is not clear is exactly in which aspects the current practice deviates from that model, and why. We believe the current debate is hindered by the fact the process leading to journal articles from conference proceedings in Computing disciplines has been largely unexplored (Montesi and Lago, 2008).

Our goal here, is to explore that process. Of course, explaining how and why Computer Sciences have attained their atypical publishing practices is a labyrinthine task, which we have no pretense to exhaust. We aim instead at at keeping a neutral point of view, and focusing on data gathering and statistical analysis. We expect that our data will help to advance the current discussions, providing a better understanding of the status quo, revealing unsuspected patterns, and grounding the intuitive ones in hard data.

2.2 Self-citation Practices

Previous research has not placed the emphasis, as we do, in self-citations as indicative of potential continuations of research lines, but in quantifying the fraction of all citations they represent. Aksnes (2003) studied self-citations in

papers published by Norway authors, discovering, among other results, that 24% of the citations in Computer Science papers were self-citations. Hyland (2003) also analyses self-citations, using a small set of articles published on leading journals in 8 disciplines. That set did not include Computer Sciences, but for the two closest disciplines represented, Mechanical and Electrical Engineering, respectively 11.8% and 9.3% of the citations were self-citations.

Other authors (Fowler and Aksnes, 2007; van Raan, 2008) are more interested in the impact of self-citations in bibliometric measures, such as the h-index and the total citation count. Finally, some authors (Frandsen, 2007) are concerned with the practice of self-citation at the journal level, i.e. the practice (maybe encouraged by publishers) of papers in a journal citing other papers in the same journal, to inflate its impact factor.

Although comparative publishing practices (including citation practices) among disciplines has been a topic of interest (Lisée et al., 2008; Montesi and Lago, 2008), the continuation of research lines is a novel angle to those investigations. We are unaware of previous studies in any scientific discipline, except for the work of Eckmann et al. (2012), which also relates to Computer Sciences. That study bears some similarities to this one, but their focus is the relationship between papers accepted in top rated Computer Vision conferences, which were later published on top ranked journals. The aim was to understand how much the publication in a top conference was a requirement to publishing on the top rated journals, both as indicated by the data, and as explicitly expressed by the community opinion. Although the emphasis was not on the continuation of a research line, clearly the conference published papers were either extended or continued in the later journal publications. The study proposed here aims both at a broader scope (all of Computer Science, instead of Computer Vision) and more general questions.

Duplicate publication is a tangential issue to our work: our aim is not to analyze republications, but we were aware of the phenomenon, since some of the self-citations turned out to fall in that category, especially on the conference originals sample. Studies on duplicate publication can be found mainly for Medical Sciences (von Elm et al., 2004), but there is one study focusing on technological areas (Kostoff et al., 2006). In addition to the obvious ethical issues, the Medical literature is particularly interested on how duplicate publication of the *same experiment data* might bias systematic reviews and meta-analyses of a particular topic.

3 Data and Methods

We have selected a random sample of 100 papers published on ACM sponsored conferences and 100 papers in ACM sponsored journals, all in the year of 2003. The data collection took place in October 2008 for the conference papers, and in November 2011 for the journal papers, both from the ACM Digital Library (ACM DL). The aim in both samples was selecting full, peer-reviewed, scientific articles, representative of a large array of CS disciplines. Therefore, articles with

4 pages and less were removed, to avoid considering extended abstracts, short papers (published on conferences) and editorials (published on journals). For the journals sample, we have removed both the ACM newsletters (for example ACM SIGPLAN Notices) and the magazines (for example Communications or Interactions), which have many regular columns and invited articles. For the journal sample, we have excluded two atypical cases: the very large volume 22 issue 3 from TOG which published the 2003 SIGGRAPH proceedings, and the JACM volume 50 issue 1, which published a commemorative issue for 50 years of the JACM, with many invited articles. We have called the set of 100 conference papers the **conference originals** and the set of 100 journal papers the **journal originals**. Collectively they are know as **originals**.

We collected the citations to the originals, also using the ACM DL, in November of 2011. The papers that cite the originals, and that were co-authored by any of the original's authors were considered follow-ups. All papers were took into account for the follow-up analysis, including short ones — provided they were more than a simple abstract, in order to allow the analysis. Note also that the follow ups come from "The Guide to Computing Literature" of ACM DL, and thus, include journals and conferences from CS at large. We have then analyzed each follow-up to verify if indeed they were true continuations of the originals. After examining a few dozen samples, we devised a three-way classification of the self-citations for the follow-ups:

- **state of the art:** the original paper was listed as part of the previous art to the follow-up paper, but was not otherwise related to it. In those cases, the original work was cited among others in the related research section of the follow-up, with or without any further elaboration. For example the citation to Eckmann et al. (2012) in the previous section is a state-of-the-art self-citation. That paper was used in the related literature section to place this research in the context of others, but this research does not *continues* that one, although they are both in the general area of conference and journal publications in CS.

- **extension:** if the follow-up paper uses or extends the original paper.

- **republication:** if the follow-up is the same paper as the original, allowing for formatting and minor text differences. It also includes papers with deletions from the original, but no new material.

We have analyzed the papers by looking mainly at the context of the citations to the originals and the papers' abstracts, but we have often looked elsewhere in the fulltext, when those evidences were not sufficient to make a decision. Each of us, the authors, have analyzed the papers independently, and where we have disagreed on the classification, we have discussed our points of view, until a consensus was reached. The disagreements happened in a minority of cases, and normally we could reach consensus quickly, pointing out at some evidence the other author had overlooked. Just a handful of borderline cases have taken, each, a few dozen minutes to conciliate.

Our classification tries to represent a spectrum of relatedness that goes from weakly related (state of the art) to almost exact reproduction (republication). However, we have only considered that the a follow-up was a true continuation if the self-citation falls in the extension class. In that case, we will say that the follow-up was a true continuation or an **extension** of the research line represented by the original publication.

For all extensions of the originals, we have also collected data regarding the time of publication, the number of authors in common with the original, and whether the continuing authors were junior or senior researchers. The status of an author as junior or senior was determined by looking at the author data in the ACM DL. If the author had any paper listed with publication date on or before 1998 (5 years before the publication of the originals) he or she was considered a senior researcher, otherwise, a junior one. Since a PhD usually lasts for around 5 years, and PhD students are the ones we would like to consider junior researchers, that interval was used to distinguish them from their advisors, considering unlikely that a typical PhD student would have publications before their graduate study years. That collection of junior/senior data was performed in November 2011.

We have also collected data on the "quality" or selectiveness of the original conferences and journals. The quality of the conference was determined by verifying the publication data about the conference in the ACM DL. From the list of all ACM proceedings[1], links to the specific conferences provide, in most cases, the total number of citations, the total number of papers accepted, and the acceptance rate. When there was no data for the acceptance rate on the year 2003, we have used the data for the earliest date available after 2003. For 8 of the conferences represented in the sample, no data was available at all, so we have removed those conference from the calculations related to "quality." The metrics of quality employed are both the rejection rate and the average number of citations received by the papers published on the conference that year. Rejection rate is an *a priori* measure of quality of a conference. One expects that conferences with higher rejection rates are selecting better papers to be published, although Freyne et al. (2010) have recently shown that for the same conference the rejection rates does not correlates with higher citations. Citations per paper is an *a posteriori* measure. If the conference had indeed "better" papers, they should receive, on average, more citations.

The "quality" of the journals was determined by using Thompson Reuters' journal impact factors, as published on the 2003 Journal Citation Records report. Unfortunately, of the 19 different ACM journal represented in our sample, 8 had no Thompson Reuters' impact factors in 2003. As with conferences, we removed those 8 journals from the calculations.

Finally, the acknowledgment of the research by the CS community was measured by whether the paper had received at least one citation that was not a self-citation, i.e., from a paper that did not include any of the original co-authors of the paper. Although not perfect, we have used that as a metric

[1] `http://dl.acm.org/proceedings.cfm`

that the paper was "accepted" outside its original group.

3.1 Statistical Analysis

Our goal is to estimate certain *parameters* of the entire publication population from our sample. The standard way is to report confidence intervals for those estimations, reflecting interval of values one expects (with a certain confidence) the population parameter will fall, given the measures taken in the sample. Employing the confidence of 90%, we compute the confidence intervals respectively for proportions, means and medians using an adapted binomial test (implemented as `prop.test` in the statistical package R), the standard t-test (`t.test` in R), and a bootstrapping procedure with 10,000 repetitions (`boot.ci` in R). We abbreviate the 90% confidence interval as $(CI = a, b)$, where a and b are, respectively, the lower and upper limit of the interval. For the correlation studies we employ a linear regression, considering models and parameters significant when their p-value is below 0.10 (reflecting, again 90% confidence).

4 Results

We start by presenting the results of our survey, without any commentary beyond the statistical analysis. We invite the reader to think about these data and their meaning, before looking at our own interpretations and conclusions in the next section.

There were 7134 proceedings articles published by ACM in 2003, which roughly correspond to our population of conference originals. We say roughly because the original set contained papers with less than 4 pages which were excluded from our sample. From the surviving ones, we have sampled 100. There were 533 journal papers in 2003, which again corresponds roughly to the population, except that this set contains the articles published on the excluded TOG and JACM issued, besides the papers with less than 4 pages, which were excluded, as before. Again, from the surviving ones, we have sampled 100. There were 20 different journals published in 2003, from which 19 appeared in the sample; and there were 147 different conferences, from which 44 turned out in the sample.

4.1 Follow-ups

A significant fraction of papers — about half — does not even have a follow-up. From the 100 conference papers sample, 44 have no follow-ups; the fraction is even larger for journal papers: 51 from the 100 samples. A small number of conference had several follow-ups, up to 14 for one of the originals. That did not happen for the journal papers: the follow-ups, if present at all, were at most one per original. The entire distributions for both samples can be seen at Table 1.

For the 100 conference originals, the 127 follow-ups, were divided evenly between journal (63) and conference (64) publications. For the journal originals, these 49 follow-ups were divided into 28 journal papers, and 21 conference papers (3 of which appeared in ACM newsletters that published conference proceedings, being thus considered conference papers).

Table 1: Follow-ups for conference and journal originals. "90% CI" indicates the 90% confidence interval on the number of originals.

	Conference		Journal	
Follow ups	Number of originals	90% CI	Number of originals	90% CI
0	44	36-52	51	43-59
1	28	21-36	49	41-57
2	11	7-11		
3	7	4-12		
4	5	2-10		
5	3	1-7		
6	1	0-4		
> 6	1	0-4		

4.2 Extensions

The analysis of the 127 follow-ups for the conference originals, and the 49 follow-ups for the journal originals, revealed, as expected, that not all self-citations are examples of research continuation. In particular, there were 57 cases of state-of-the-art and 15 cases of republication citations among the follow-ups of conference originals (28 and 0, respectively, for the follow-ups of journal originals). For two of the follow-ups of journals, it was impossible to access the fulltext in order to evaluate them.

Table 4.2 lists the extension counts for the journal and conference originals. The 52 extensions of the conference originals were divided into 23 that were published on journals, and 29 in conferences. The 21 extensions for the journal originals were divided into 14 journal papers, and 7 conference papers (including 3 published on newsletters as discussed above).

4.3 Lapse between original and extension

The time interval between the originals and extensions is shown in Figure 1, in the form of a boxplot for each population of originals. The boxplot showcases the medians, quartiles and range of the populations. In addition, we have computed the averages and confidence intervals. For originals published on conferences the average interval was 1.7 years ($CI = 1.3, 2.2$). Breaking up this average further, it takes 1.4 years ($CI = 0.7, 2, 2$) if the extension is published on a journal (9

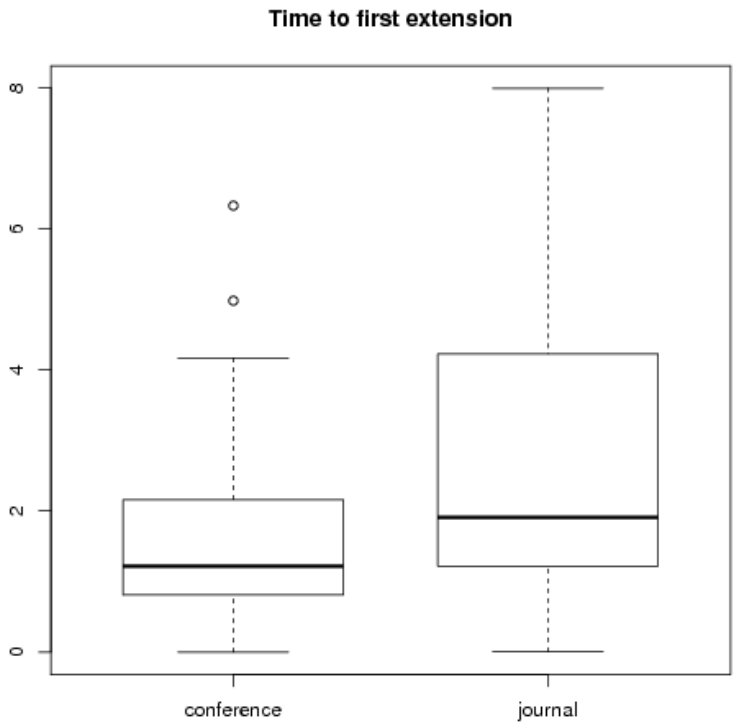

Figure 1: Boxplot of the time interval for the first extension, for conference and originals

.

Table 2: Extensions of conference and journal originals

	Conference		Journal	
Extensions	Number of originals	90% CI	Number of originals	90% CI
0	70	62-77	77	71-85
1	16	11-23	21	15-29
2	9	5-15		
3	3	1-7		
4	1	0-4		
5	1	0-4		

cases), and 1.8 years ($CI = 1.2, 2.4$) for a first extension to be published on a conference. For journal originals, the average time for the first extension was 3.0 years, ($CI = 2.0, 3.9$); with 3.2 years ($CI = 1.2, 5.2$) to publish it on a conference, and 2.8 years ($CI = 1.7, 4.0$) to publish it on a journal.

4.4 Who continues the work

For conferences, on the average, 62% ($CI = 55, 68$) of the authors of the original were also authors of the continuation. When we consider only the senior researchers in the originals, 65% ($CI = 56, 73$) of them continue as authors in the extension. For junior authors, 58% ($CI = 48, 67$) of them remain as authors in the extension. There were 24 extensions in which only one of the original authors remained, and in these cases, 10 were extended exclusively by the junior author.

For journals, 55% ($CI = 44, 65$) of the authors remained as co-authors of the extensions. The continuation rates for senior authors was 62% ($CI = 49, 74$) and for juniors 37% ($CI = 19, 58$). Of the 10 cases in which only one author remained, only 1 of them was extended by the junior author.

4.5 Acknowledgement by community *vs.* extension by authors

We contrast four populations of originals in the boxplot of Figure 2: conference and journal originals, subdivided further into those that where extended, and those that were not. For each population, we plot the number of citations (excluding self-citations). Outliers, which were twice the interquartile range from the median, are indicated by small dots in the boxplot. In addition, some extreme outliers not shown in the figure, to avoid compressing too much the scale: 188 citations to the conference originals with extension; 439, 259, 134, and 130 citations to journal originals without continuations; and 130 citations to journal originals with continuation.

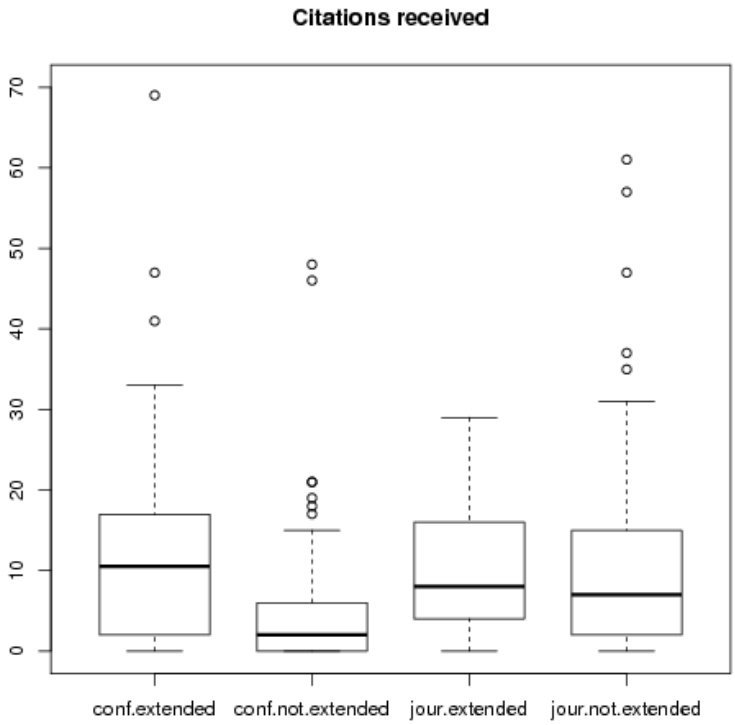

Figure 2: Boxplot of the citations received by conference originals with and without extensions, and journal originals with and without extensions. There is one outliers of 188 citations for conference originals with extension, four outliers of 439, 259, 134, and 130 for journal originals without citations, and one outlier of 130 citations for journals with continuations not shown in the figure.

The median citation count for conference papers with extensions is 10.5 ($CI = 6, 13.5$), while for papers without extension, the median is 2 ($CI = 1, 2$). Thus, for the conference population, the difference appears significant, both visually in the plot and statistically.

For journal originals, the median citations for papers with extensions is 8 ($CI = 4, 13$), and for papers without extension, the median is 7 ($CI = 4, 7$). Both the plot and the statistical analysis seem to indicate that there is not much difference.

For the conference originals, 3 papers out of the 30 with extension had no citations other than self-citations, something that happened for 20 out of the 70 papers without extension. For the journal originals, the numbers are 1 out of 21 for the papers with extension, and 8 out of 77 for the others.

4.6 Quality of the medium

For each of the samples, we have investigated the correlation between the proportion papers that were extended against some measure of quality or selectivity of the conference or journal, accordingly. For the conferences we have employed the acceptance rate and the citation count per paper (for all papers in the conference, not just the ones in our sample) as quality metrics. For the journals, we have used the impact factor.

In none of the cases, the regression analysis could find a significant influence between the quality metric and the proportion of extensions (the p-values found were typically below 0.60). Visual inspection of the data plots confirmed that lack of influence.

5 Conclusion and discussion

We can now address the questions put forth in the Introduction of this paper.

Part of the answers come straightforwardly from the data, like how many papers do effectively continue. We find out that only 23% (journal) and 30% (conferences) of the papers published in CS have extensions, indicating a continued line of research. Those numbers are surprisingly low, but as we have mentioned, we lack similar results on other scientific disciplines to properly evaluate those proportions. The difference seems to disfavor journal articles, but it is not significant at 90% confidence level. However, a significant difference is found in the potential *number* of extensions, since, when they do continue, conference papers might generate more continuations than journal papers.

Leaving the realm of hard data to that of interpretation, we conceive two explanations for that difference. One, more traditional, is that conferences are being used for work in earlier stages, thus, prone to more ulterior development. However, since there is no significant difference in the *proportion* of the extensions, just in their *quantity*, we tend to believe that a better explanation might be found considering both the contingencies of conference publishing and the habits of the CS community. The former, due to stringent deadlines

and paper length limitations, may stimulate the research being cut into "least publishable units". The latter may make a CS researcher uncomfortable in publishing a journal paper before some of the obvious extensions of the basic idea have been duly explored in that paper. CS researchers usually favor long and complete journal papers, while conference papers are almost always shorter. Thus a journal paper would already exhaust some of the possible extensions avenues.

That hypothesis would also explain the longer time it takes for a journal original to have an extension published. As observed, the elapsed time for the publication of the extension is much shorter for conference originals, regardless if the extension is in a conference or in a journal.

Conferences tend slightly more to be extend in conferences than in journals (odds of 1.2 to 1), while journals are much more prone to being extended to other journals (odds of 2 to 1). That hints at the directionality of the traditional evolutionary model, which has not been completely abolished, even in the CS community.

For conferences and journal papers, 62% and 55%, respectively, of the original authors remain as co-authors of the extensions. When researchers do drop-off, the ones who remain tend to be the senior ones. That fact, which somewhat surprised us, suggests that the research lines are "owned" by the senior researchers, who develop it in cooperation with other faculty members and students. It is rarer for the opposite situation to happen, that is, the student to "carry away" the research line with him, after graduation.

Does the chance of continuation correlates with the community reception of the paper? The correlation is significant for conference papers, but not for journal papers. Of course, correlation does not imply causation, but one can argue that researchers in some level gauge the reception to their conference papers on deciding on whether to continue that line of research. A similar correlation was found by Eckmann et al. (2012) — papers published on top Computer Vision conferences and then, later, on a top journals, had more citations than papers from the same conferences that were not extended to journals. It is less clear why that is not observed on journal papers — one can speculate that the publication time and the time to accumulate citations are longer for journals, added to the fact that one seldom gets feedback from the peers for journal papers. Thus the validation, or lack thereof, provided by journals may not be timely enough to guide the researcher in his decision to continue or not the research line.

More surprising is the lack of correlation between the quality or selectivity of both journals and conferences and the amount of papers that continue. The correlation observed for conference papers in terms of *individual* impact is not observable in terms of the *aggregate* metric provided by either the conference selectivity or its collective impact factor. We were expecting either a positive correlation (meaning that the acceptance stimulates the authors to further pursue the research), either a negative correlation (meaning that the publication crowns the research line, liberating the authors to other pursuits). Is it possible that both effects happen in balance, creating a net effect of no correlation?

Finally, let us discuss the 70–77% of papers that have no continuations. Do they reflect an accomplished result, the terminus of a fruitful research line, or a speculative attempt that did not pan off and was abandoned? Again, we enter the realm of interpretation, and the criteria we elected is not meant to be waterproof. We consider the research line "accomplished" if that not-continued paper has at least one citation by papers outside the original group, i.e., if it was considered relevant enough by outsiders to be cited at least once. That criteria is, of course, arbitrary, and at once too stringent (a paper can be inspiring and end up not being cited) and permissive (a paper can be cited to be criticized or rebuked). Yet, though we feel the "one outside citation" criteria is useless to judge the merit of any individual paper, we believe it is useful as an *statistical* tool, since the extreme cases get smoothed out.

Considering that, the figures indicate that conferences are indeed, even for Computer Sciences, used for more speculative ideas, with 29% ($CI = 20, 39$) of papers (20 out of 70) without citations other than self-citations. That proportion is much smaller for journal papers, 10% ($CI = 5, 18$) of papers (8 out of 77) in that situation. Again, that hints at the presence of the evolutionary model, even if diluted, in CS.

That presence is also suggested by a small aside of our research, about the works which were *republications* of the original, as explained in Section 3. A non-negligible fraction of authors follows the standard practices of other disciplines, by publishing first in a conference and then following-up on periodicals: 26% ($CI = 19, 34$) of all conference originals were either extended or republished on journals.

References

Aksnes, D. (2003). A macro study of self-citation. *Scientometrics*, 56(2):235–246.

Birman, K. and Schneider, F. (2009). Program committee overload in systems. *Communications of the ACM*, 52(5):34–37.

Eckmann, M., Rocha, A., and Wainer, J. (2012). Relationship between high-quality journals and conferences in computer vision. *Scientometrics*, 90(2):617–630.

Fortnow, L. (2009). Time for computer science to grow up. *Communications of the ACM*, 52(8):33–35.

Fowler, J. and Aksnes, D. (2007). Does self-citation pay? *Scientometrics*, 72(3):427–437.

Frandsen, T. (2007). Journal self-citations: Analysing the JIF mechanism. *Journal of Informetrics*, 1(1):47–58.

Freyne, J., Coyle, L., Smyth, B., and Cunningham, P. (2010). Relative status of journal and conference publications in computer science. *Communications of the ACM*, 53(11):124–132.

Grudin, J. (2011). Technology, conferences, and community. *Communications of the ACM*, 54(2):41–43.

Halpern, J. and Parkes, D. (2011). Journals for certification, conferences for rapid dissemination. *Communications of the ACM*, 54(8):36–38.

Hyland, K. (2003). Self-citation and self-reference: Credibility and promotion in academic publication. *Journal of the American Society for Information Science and Technology*, 54(3):251–259.

Kostoff, R., Johnson, D., Antonio Del Rio, J., Bloomfield, L., Shlesinger, M., Malpohl, G., and Cortes, H. (2006). Duplicate publication and "paper inflation" in the fractals literature. *Science and Engineering Ethics*, 12(3):543–554.

Lisée, C., Larivière, V., and Archambault, É. (2008). Conference proceedings as a source of scientific information: A bibliometric analysis. *Journal of the American Society for Information Science and Technology*, 59(11):1776–1784.

Montesi, M. and Lago, P. (2008). Software engineering article types: An analysis of the literature. *Journal of Systems and Software*, 81(10):1694–1714.

van Raan, A. (2008). Self-citation as an impact-reinforcing mechanism in the science system. *Journal of the American Society for Information Science and Technology*, 59(10):1631–1643.

von Elm, E., Poglia, G., Walder, B., and Tramer, M. R. (2004). Different patterns of duplicate publication an analysis of articles used in systematic reviews. *JAMA*, 291(8):974–980.

Wing, J. and Guzdial, M. (2009). CS woes: deadline-driven research, academic inequality. *Communications of the ACM*, 52(12):8–9.