



DONALD E. KNUTH
Professor Emeritus of The Art of
Computer Programming
Department of Computer Science
Telephone [650] 723-4367

October 25, 2003

Editorial Board, *Journal of Algorithms*

Dear Board member,

I hope you don't mind receiving this somewhat extraordinary letter—and indeed, I confess it's also quite a long one, because I've never learned how to be brief. The growing turmoil in the world of scholarly publishing has been weighing heavily on my mind for several years, and I've finally decided to ask for your collective wisdom on what future direction our beloved journal should take.

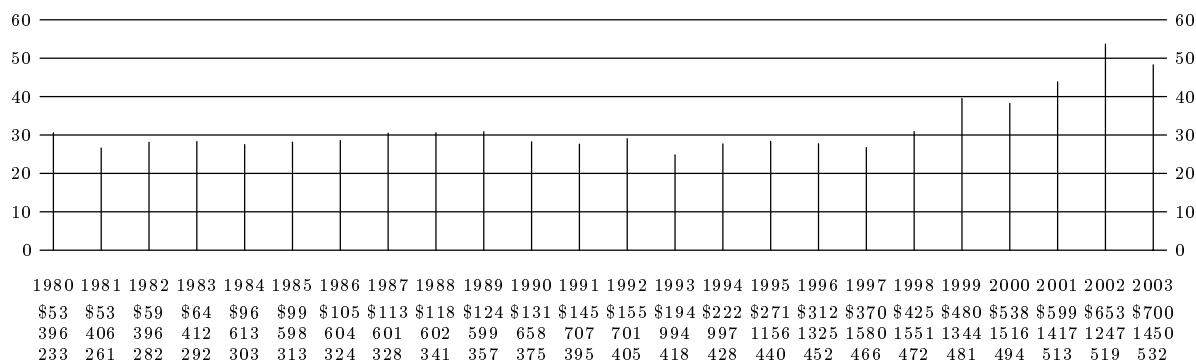
Let me begin with some background information from my personal perspective. I “grew up” professionally with Academic Press journals: Part of my thesis was printed in Volume 2 of the *Journal of Algebra* (1965); soon afterward I published an article about trees in the *Journal of Combinatorial Theory*, Volume 3. I was eventually destined to publish six more papers in the latter journal, and one each in the *Journal of Mathematical Analysis and Applications*, the *Journal of Number Theory*, and the *Journal of Computer and System Sciences*. Those papers were typeset so beautifully, I used Academic Press style as the model in my first demo of \TeX to the American Math Society in 1978.

Therefore I was pleased when Herb Wilf approached me later that year with the idea to start a new Academic Press publication, to be called the *Journal of Algorithms*. On January 4, 1979, I replied to him that “*Journal of Algorithms* is a great title. Surely there must be a journal of that name someday.” We agreed that computer science had matured to the point where such a journal would be an ideal outlet for some of the explosive growth in high-quality algorithmic research.

Over the years the issues of this journal have accumulated to fill nearly five feet of shelf space in my office at home, and I couldn't be more proud of the quality of many of the articles they contain. The experience of compiling and typesetting the index to Volumes 1–20 that appeared on pages 634–660 of the May 1996 issue gave me a special pleasure; and next year we shall reach Volume 50.

Academic Press built its reputation on producing high-quality scientific books and journals at reasonable prices. That is why Wilf and I were attracted to them initially, and why we continued to be satisfied as the years went by. Academic Press was acquired in 1989 by Harcourt Brace Javonovich, later to become known as Harcourt X for various other values of X, but at first their publishing team stayed fairly intact. I became concerned about journal pricing in 1990, and I wrote a two-page letter asking them to do their best to minimize the effect on libraries; they promptly sent me a completely satisfactory reply, and indeed they kept price increases below the level of inflation during the next few years.

That policy changed rather abruptly at the end of the 90s, when the cost of the *Journal of Algorithms* began to increase substantially. Then, in May 2001, Reed Elsevier purchased Harcourt General for \$4.5 billion, and the price of *JoA* has continued to increase in spite of my protests. Here is a chart showing the actual figures in cents per page, converted to 2003 dollars in order to offset the effects of inflation:



(The three figures shown below each year are (1) the cost of a library subscription; (2) the number of pages; and (3) the consumer price index in January, taken from <ftp://ftp.bls.gov/pub/special.requests/cpi/cpia1.txt>. The number of pages in 2003 is, of course, an estimate; I computed it by taking 4/3 of the pages in the first nine months and multiplying by 1.4 (since this year's pages each contain 40% more text than last year's). In summary, the cost in constant dollars held steady for many years between about 25¢ and 30¢ per page, until shooting up past 50¢ per page in 2002.

Of course you all know that the publishing environment has changed dramatically since 1980. In those days a publisher took care of keyboarding and proofreading, which were quite expensive, especially when mathematics was involved. Only a few printing establishments in the world were able to do a good job of that, and we gladly paid for the privilege of having our work presented well. But now, authors have taken over most of that work, and software has also ameliorated the other aspects of a publisher's task.

So the cost per page should have dropped. Instead, it has doubled. The same has happened with virtually every other mathematics journal produced by commercial publishers, often with ratios far in excess of 2 to 1. Journals produced by non-profit organizations have generally kept costs steady, but the for-profit publishers now control about 60% of the leading journals.

I love my library and the other libraries that I visit frequently, and my blood boils when I see a library being overcharged. Therefore I wrote a strong letter to Elsevier in August 2001, just after I learned that they had taken over Harcourt, expressing serious concerns about their future pricing policy for the *Journal of Algorithms*. (At that time I was also incensed about the fact that Elsevier was charging Stanford \$4878 for a year's subscription to *Theoretical Computer Science* and *Science of Computer Programming*.) Elsevier, however, ignored my letter and did not reply.

Meanwhile I had learned that many other people and organizations had been getting together and taking action to reverse these alarming trends. In the next part of this letter, I'll quote extensively from some of the best discussions that I've encountered in the literature, giving also Internet citations so that you can read further; I encourage you to do so, because these sources are packed with interesting facts.

First, I strongly recommend the article "Scientific publishing: A mathematician's viewpoint" by Joan Birman of Columbia University, which appeared in *Notices of the AMS* **47**, 7 (August 2000), 770–774. Assuming that your institution subscribes to American Math Society publications, you can access this article online at www.ams.org/notices by clicking on "Forum" when you get to the relevant issue.

Joan began by saying “We are in a time of ferment with regard to the ways in which mathematical research is being communicated throughout the world, and in particular with regard to the nature and cost of scientific journals. . . . The publishers of some of our best journals . . . have begun to charge such high prices for library subscriptions that to continue with them means to threaten the rest of the collection, but to drop them means to create a big hole in the collection. . . . If the people who do the research (the leaders and future leaders in the field) are prepared to act, then the entire system can be changed and the problems solved. . . .

“[We know] that the essential value in a journal article comes from the excellence of the work of the author and the value added by mathematical colleagues. . . . We have learned to ‘typeset’ our own papers, do our own graphics, and in general deliver beautiful manuscripts that are printer-ready when we submit our papers. . . . When the moment comes to choose a journal, others in the mathematical community have already contributed, via their expertise, to the reputation of the journal. . . . Papers are refereed in a careful and serious way. . . . Finding a referee is minor compared to doing the job, which can be hard work. . . . The refereeing process adds value to the paper, not in an easily quantifiable way. . . .

“The moment when a paper is accepted is also the moment when the author is asked to sign a copyright agreement. Since profits, if any, go to the owner of the journal, I wondered whether dollar costs had been incurred by the owner of the journal at any point up to this moment? From conversations with many colleagues, I verified that essentially all the work described above is done pro bono, with the possible exception of the contributions of the journal editors.”

She mentioned a case where the editor-in-chief of the transactions of a professional society received \$12,000 for secretarial support at his department. She heard rumors that editors-in-chief were essentially being bribed by publishers so that they would have no incentive to argue about pricing; but two editors vigorously denied this claim. She learned that the annual salaries for editors of four journals were respectively \$6K, \$12K, \$14K, and \$22.5K. I myself receive \$1,000 per year from Elsevier, in my role as one of the founding editors, plus \$1,667 for secretarial support. I also receive two complimentary copies of the *Journal of Algorithms*, one of which goes to the Theory Research Collection used by Stanford’s grad students and faculty, located near our offices.

Members of the *JoA* editorial board at large receive free copies, and I know that this is not insignificant: For example, I receive gratis copies of the *Journal of Computer and System Sciences*, another journal that I believe to be overpriced—the cost per page in 2000 was 86¢, which would be 65¢ per old *JoA* page because their pages are about 33% larger than ours used to be. I suppose I should have resigned from their editorial board as a matter of principle, but in fact those copies are important for me as I work on *The Art of Computer Programming*. (The same holds for *Acta Informatica*, published by Springer at a cost to libraries of \$1.82(!) per old *JoA* page in 2002; *Random Structures & Algorithms*, published by Wiley at 64¢; *Software Practice & Experience*, also Wiley, at \$1.26.) The more a journal costs, the more I need to be on its editorial board.

Another perk available to me as one of the main editors of the *Journal of Algorithms* is total access to all of Elsevier’s ScienceDirect electronic library, a privilege available at the moment to people at only a few universities like Ohio State, where a deal was negotiated long ago at “introductory rates.” But I have not used that privilege; somehow I would feel guilty doing so.

Birman went on to make a useful classification into journals of four types:

- (i) owned by universities, e.g. the *Annals of Mathematics* (Princeton and IAS) and the *Pacific Journal of Mathematics* (produced by a consortium of 34 universities including several in China and Japan);
- (ii) owned by a learned society like ACM or SIAM;
- (iii) owned by a university press, like Cambridge or Oxford or Duke, etc.—sometimes hard to distinguish from type (i);

(iv) owned by a commercial publisher.

She estimated that roughly 17%, 13%, 10%, and 60% of the best math journals belong to these respective categories.

Rob Kirby has compiled extensive data on the costs of 128 math journals (including *JoA*) between 1995 and 1999; see math.berkeley.edu/~kirby/jp00.html. His list makes it abundantly clear that the type (iii) and (iv) journals are significantly more expensive than the others. For example, the *Annals* and the *Pacific Journal* weigh in at about 11¢ per page (and those costs have subsequently decreased); *SIAM Journal on Applied Mathematics* at about 16¢; *Combinatorics, Probability, and Computing* (Cambridge University Press) at about 40¢; and there are ten different Springer journals that cost about \$1.00 per page. To supplement his data I can report that *JACM* costs less than 20¢ per page, and those pages hold roughly 5/4 as much as the old *JoA* pages used for comparison in my chart. Various ACM transactions have different costs depending on the subject matter: *TOMS* was 21¢ per page in 2000 (but 33¢ in 2002!); *TOPLAS* has been 18¢, 19¢, and 19¢ during the past three years; and the *ACM Transactions on Graphics* went from 41¢ to 57¢ to 19¢ for some reason. *SIAM Journal on Computing* checks in at 19¢, 20¢, and 22¢; but *SIAM Journal on Discrete Mathematics* is way off the curve, namely 56¢, 60¢, and 62¢ for 2000, 2001, and 2002.

Kirby had made a previous survey in 1997, where he also figured out the cost per 10000 characters, as of 1996 (math.berkeley.edu/~kirby/journals.html). In this list he seriously undercounted the number of pages in the Elsevier journal *Nonlinear Analysis*, making those estimates too high by a factor of 2.4; this anomaly was corrected in his other tabulation cited above. If we adjust this erroneous data point, and if we also discount one journal that is translated from Russian at an obviously greater expense, we find that the venerable publishing house of Springer Verlag emerges strongly as the main culprit among type (iv) publishers, with journals whose “scores” are respectively 357, 348, 335, 320, 318, 311, 306, 272, and 244 cents per 10000 characters. Wiley’s *Communications on Pure and Applied Mathematics*, at 276, was Springer’s only “competitor.” Elsevier’s most expensive journals according to his list (which did not include *Theoretical Computer Science*) were: *Applied Mathematics and Computation* (238); *Topology and Its Applications* (209); and *Discrete Mathematics* (200).

In 1996, the *Journal of Algorithms* (74) turned out to be the most economical of all Academic Press journals in Kirby’s ranking, while the least expensive from Elsevier was the *Journal of Computational and Applied Mathematics* (108). Curiously, *Algorithmica* (52) was significantly less expensive than *JoA* by Kirby’s measure, proving that Springer journals are not always high-priced; the only other Springer journal rated less than 150 was, however, *Semigroup Forum* (89). (The cost of *Algorithmica* in 2002 was $856/1945 \times .8 = 35¢$ per old *JoA* page.) If we look at the 37 journals that Kirby rated less costly than the *Journal of Algorithms* in 1996, it turns out that *Algorithmica* was the only journal of type (iv), and there were only seven of type (iii). The ten least expensive, including the *Pacific Journal*, were rated 39, 37, 34, 33, 32, 29, 28, 25, 23, and 20—an order of magnitude less than those at the high end.

But I want to return to Birman’s article from 2000, because on page 772 she proceeded to tell the interesting story of the *Journal of Logic Programming*. After more than a year of unsuccessful negotiations with Elsevier about price reductions, all 50 members of that journal’s editorial board resigned. They founded a new journal called *Theory and Practice of Logic Programming*, published by Cambridge University Press, thereby moving from category (iv) to category (iii). The *Journal of Logic Programming* ceased publication with (an abbreviated) Volume 46 in the fall of 2000. Elsevier then began its own new journal, entitled the *Journal of Logic and Algebraic Programming*, but I cannot tell you anything about it because Stanford does not subscribe. Meanwhile the new *Theory and Practice of Logic Programming* appears to be thriving. The cost to Stanford’s library in 2002 was $\$362/850 \approx 43¢$ per page, however, so I cannot call this a totally unqualified success.

Krzysztof Apt, one of the principal figures in the conversion from *JLP* to *TPLP*, published an excellent article entitled “One more revolution to make: Free scientific publishing,” in *Communications of the ACM* **44**, 5 (May 2001), 25–28. “Thanks to the Internet,” he wrote, “computer scientists are in a position to create new, free high-quality journals and contribute in this way to a free dissemination of scientific knowledge. . . . You don’t need to be a genius to realize that scientific publishing is a very peculiar branch of business . . . it is the only branch of industry that relies on massive voluntary work.” He went on to give specifics, concluding that computer scientists were lagging behind other fields—observing, for example, that more than 11,000 scientists from biomedicine had signed an open letter pledging that “beginning in September 2001, we will publish in, edit or review for, and personally subscribe to, only those scholarly and scientific journals that have agreed to grant unrestricted free distribution rights to any and all original research reports they have published.”

See also Apt’s article “Towards free access to scientific literature,” *Nieuw Archief voor Wiskunde* (5) **2** (2001), 251–255 [homepages.cwi.nl/~apt/ps/naw.ps], which contains pertinent remarks about how commercial publishers are currently disguising their pricing policies by bundling several different journals into packages, signing individual contracts with libraries that bind them to confidential multiyear subscriptions at fixed annual rate increases. A clever ploy, with no accountability because the rates are not disclosed. “Deals offered by learned societies don’t have such offensive clauses.”

The *Journal of Machine Learning and Research* was founded in 2000 by 40 of the 67 editorial board members of *Machine Learning*, a Kluwer journal that cost more than \$1000 at the time. Those editors formed JMLR, Inc., which offers a free full-text electronic edition. There’s also a paid electronic edition, which includes additional links to abstracting, indexing, archiving, and mirror sites; and a hardcopy edition published quarterly by MIT Press. [See www.jmlr.org/statement.html, where Michael Jordan wrote, “[Fees charged by publishers] provide access for institutions and individuals who can afford them, [but] we feel that they also have the effect of limiting contact between the current machine learning community and the potentially much larger community of researchers worldwide whose participation in our field should be the fruit of the modern Internet. None of the revenue stream from the journal makes its way back to authors, and in this context authors should expect a particularly favorable return on their intellectual contribution—they should expect a service that maximizes the distribution of their work.”] JMLR is currently experiencing a cash shortage, however, and I’m told that MIT Press is unhappy with the present arrangement. The original journal, *Machine Learning*, has meanwhile come back and its latest issue lists 59 members of the editorial board. Stanford’s library has not saved any money in this case, nor has it lost any (because it subscribes only to the online edition of *JMLR*). On the other hand, *JMLR* had the highest “ISI impact factor” of all journals in Artificial Intelligence in 2002; thus its future does not really look bleak, in view of recent developments in online publishing that I shall discuss below.

But let’s return to Joan Birman’s article: She went on to describe a successful “type (i)” journal from the late 90s, *Geometry and Topology*, hosted by the math department at Warwick University. She quotes co-founder Colin Rourke’s observations that “Computer costs are negligible, given the fact that universities are already networked . . . I estimate that the size of the Warwick Maths computer system is about four orders of magnitude greater than that need to run *G&T*. But then this is the whole point: Journals are firmly based in the academic world, and all piggy-back to a great extent on that world. . . . Most of [the result of ten weeks of really hard work that was needed to get going] is replicable, and we could very quickly set up another similar journal or help others to do so.” Needless to say, all of the journals I’m discussing have a world-class editorial board and high stands of excellence for accepting papers; three Fields medalists and “a long list of other very distinguished mathematicians” serve *G&T*. The online version is free, and a hardcopy version is printed at the University of Warwick and distributed at an extremely reasonable price; for example, Stanford paid \$114 for Volume 5 in 2001, totalling 945 pages (thus 12¢ per page).

Geometry and Topology now has a particularly noteworthy feature that hasn't yet spread to many other journals, but I expect it soon will: All of its papers are permanently archived at Cornell University's "arXiv," formerly (or perhaps concurrently) based at the Los Alamos National Laboratory. Please go now with your browser to www.maths.warwick.ac.uk/gt/, then click on Volume 7 (2003), and then on the first article (which is about elementary knot theory). On this page you'll see a conventional abstract, plus links to the text in four versions and even to a reprint cover, plus links to the authors' home pages and email addresses. But most impressive to me is the line "E-print: arXiv:math.QA/0204311"; from here you can get a permanently archived file, available in many other formats. For example, I downloaded `0204311.tar.gz`, which includes a variety of TeX source files together with encapsulated PostScript and even a METAFONT file that generates an interesting font for graphics.

Rourke observed, "At the very least LANL seems as secure in its future as the commercial publishers." *Geometry and Topology* is also archived at many sites as part of EMIS, the European Mathematical Information Service, a worldwide server network organized by the European Mathematical Society. "Mirrors have been installed on all continents (except Antarctica)." The EMIS electronic library presently includes 61 mathematics journals that each have been judged to meet high standards of excellence, and ten more are about to follow soon [see www.emis.de/journals]. A satellite meeting of last year's International Congress of Mathematics in Beijing was held on the general topic of mathematical journal publishing, hosted by Tsinghua University, and the proceedings appeared this year as *Lecture Notes in Computer Science* **2730**, entitled *Electronic Information and Communication in Mathematics*. Curiously Springer has elected *not* to make this volume available online at its LINK website.* But if your library has purchased the hardcopy edition, you can find a complete discussion of EMIS's state as of early 2003, in the chapter by Michael Jost and Bernd Wegner on pages 89–96.

Another electronic journal to look at is the *New York Journal of Mathematics* [nyjm.albany.edu], if you can avoid being distracted by its constantly animated logo. The same website also currently hosts the *Pacific Journal of Mathematics* mentioned above, and the prestigious *Journal of Differential Geometry*. All three journals are presented in the now-familiar format of pioneering journals like Herb Wilf's *Electronic Journal of Combinatorics* [www.combinatorics.org]; namely, they offer links to full text in either PostScript or Portable Document Format (PDF), without the E-print and other features I mentioned for *G&T*. Incidentally, the PDF files of all these online journals could be improved; they include rather useless "thumbnails" and no bookmarks. I hope style files in future will extract appropriate bookmarks from section headings, as I do in CWEB output. But the PDFs do seem to work with respect to text searching; and even Adobe rarely uses PDF bookmarks well in its own product documentation.

Near the end of Birman's article she gave a call to action. "We are currently witnessing what must be properly identified: A battle for the ownership, transfer, and dissemination of scientific information." And she went on to list various conclusions, of which the first was: "Individuals who are in a leadership position can put community interests ahead of their own interests and work seriously with their colleagues on editorial boards and with the publishers to lower prices. If that fails, it is time to go shopping for another publisher."

If you are anything like me, you shrink from battles and are skeptical of your ability to lead. However, sometimes we must come out of our cocoons and do things that are necessary, even if we don't have the ability to do such things especially well.

* Note added August, 2004: At the time I wrote this letter, the electronic version was not yet available. I learned subsequently, however, that the delay was simply due to technical problems in preparing the online files, not in any way to an "election" by Springer to postpone electronic publication. Shortly after my letter appeared, this volume was indeed just as accessible as all of its neighbors were and are.

Several new developments have occurred since her paper came out. One that I cannot help but mention, although it doesn't affect the *Journal of Algorithms* directly, is the fact that Candover and Cinven, two private equity firms, bought Wolters Kluwer in January 2003 and Bertelsmann Springer in August; the combination, to be known next year as Springer, controls 1350 journals. Writeups in economic journals that you can easily find on the Web point out that this purchase is an excellent investment because scientific publishing is so profitable. The investors, advised by Merrill Lynch, intend to develop Springer's electronic assets and then sell the company in two years or so, receiving a handsome return on their investment.

An online publication entitled *Declaring Independence* [www.arl.org/sparc/DI], produced by a librarians' group called the Scholarly Publishing and Academic Resources Coalition (SPARC), with Joe Halpern representing the SIGACT community on its board of reviewers, opens by saying "Purely commercial interests have gained sway over too many of the journals that we depend on for research information. Maximizing profits has become the controlling goal. A system that should serve us is at the mercy of corporate acquisitions and profit-oriented planners." And it concludes with the words "Your journal was created for scientific stakeholders, not stockholders."

In between, *Declaring Independence* surveys numerous issues involved in journal publishing, with checklists of things to consider and pointers to various ongoing projects. In particular it mentions Project Euclid, a partnership between Cornell's math library and Duke University Press. I looked at that website [projecteuclid.org] and was not extremely impressed: They offer a package deal whereby libraries can get electronic access (only) to ten journals including *Experimental Mathematics* and *Simon Stevin* and *Journal of Applied Mathematics* and seven others I never heard of, for an annual fee of \$1250. It troubles me that librarians evidently regard this as a bargain, when most of the electronic journals known to me are free of charge.

Still another recent development is associated with the new buzzword "open access," which is slightly analogous to "open source software." A big meeting was held in Budapest during December 2001, resulting in the so-called Budapest Open Access Initiative, nicely explained at www.earlham.edu/~peters/fos/boaifaq.htm. This initiative "seeks open access for the scientific and scholarly research texts that authors give to publishers and readers without asking for any kind of royalty as payment. . . . When authors do wish to give away their writings, their readers should not have to pay access tolls to read them. . . . What's important is relinquishing payment, not relinquishing intellectual property rights. . . . By 'open access' to this literature, we mean its free availability on the public internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. The only constraint on reproduction and distribution, and the only role for copyright in this domain, should be to give authors control over the integrity of their work and the right to be properly acknowledged and cited. . . . Open-access journals do not differ from toll-access journals in their commitment to peer review or their way of conducting it, but only in their cost-recovery model, which has no bearing on the quality of the articles they publish." [See also www.soros.org/openaccess/read.shtml.]

These concepts were refined further at a meeting last April, resulting in the so-called Bethesda statement [www.earlham.edu/~peters/fos/bethesda.htm], which defines open access thus: "An Open Access Publication is one that meets the following two conditions: (1) The author(s) and copyright holder(s) grant(s) to all users a free, irrevocable, worldwide, perpetual right of access to, and a license to copy, use, distribute, transmit and display the work publicly and to make and distribute derivative works, in any digital medium for any responsible purpose, subject to proper attribution of authorship, as well as the right to make small numbers of printed copies for their personal use. (2) A complete version of the work and all supplemental materials, including a copy of the permission as stated above, in a

suitable standard electronic format is deposited immediately upon initial publication in at least one online repository that is supported by an academic institution, scholarly society, government agency, or other well-established agency that seeks to enable open access, unrestricted distribution, interoperability, and long-term archiving. *Notes:* Open access is a property of individual works, not necessarily journals or publishers. Community standards, rather than copyright law, will continue to provide the mechanism for enforcement of proper attribution and responsible use of the published work, as they do now.”

This document also states that “Our intention is to reconvene an expanded group in a few months to draft a final set of principles that we will then seek to have formally endorsed by funding agencies, scientific societies, publishers, librarians, research institutions and individual scientists as the accepted standard for publication of peer-reviewed reports of original research in the biomedical sciences.” Between the lines, this appears to imply that responsible scientists, at least in biomedicine, should agree not to publish in or edit for any journal that does not allow open access as defined above.

Do articles in Elsevier journals qualify as “open access publications”? If you look at the Information for Authors at the end of *Journal of Algorithms* **37** (November 2000), you will find the following stipulations: “2.2. The authors agree that, if the work is accepted for publication, they will transfer the copyright in it exclusively to Academic Press, including the right of reproduction in all forms and media, whether now known or hereafter developed, and the right to include it in collections and databases. . . . 3.3. After publication, authors may post their Academic Press copyrighted material on their own servers without permission, provided that the server displays as the first line of the HTML page the following notice alerting readers to their obligations with respect to copyrighted material: *This material has been published in* <name of journal, issue number and date, page numbers>, *the only definitive repository of the content that has been certified and accepted after peer reviews. Copyright and all rights therein are retained by Academic Press. This material may not be copied or reposted without explicit permission.* The posted work must also include the Academic Press copyright notice (*Copyright © 200X by Academic Press*) and a link to IDEAL (International Digital Electronic Access Library) at <http://www.idealibrary.com>.”

That was the pre-Elsevier situation. The current rules are no longer printed directly in our journal, but they are cited implicitly and authors can find them at the gateway, authors.elsevier.com; to wit, “Elsevier does not require that authors remove from publicly accessible servers versions of their paper that differ from the version as published by Elsevier. . . . Posting of the article as published on a public server can only be done with Elsevier’s specific written permission.” Presumably we could negotiate an agreement so that such permission is automatically granted; this might conform to part (1) of the open-access definition. But it would still not meet part (2). And it would require special effort on the part of each author, to try to match all changes made by Elsevier, while the system of journals like *Geometry and Topology* makes all this posting and archiving automatic and foolproof.

Quoting further from the author gateway, “Why does Elsevier believe it needs exclusive rights? The research community needs certainty with respect to the validity, normally obtained through the peer review process, of scientific papers. The scientific record must be clear and unambiguous. Elsevier believes that by obtaining the exclusive distribution rights it will always be clear to researchers that, when they access an Elsevier site to review a paper, they are reading a final version of the paper which has been edited, peer-reviewed, and accepted for publication in an appropriate journal. If Elsevier did not obtain exclusive electronic rights, it is likely that versions of scientific papers would clutter the Internet without clarification of the scientific status of such versions. See also our information on [electronic preprints](#) for more detailed discussion on these points.”

Eh? Have you ever been in doubt about the veracity of a paper found at, say, the *Electronic Journal of Combinatorics*?

Now here are some quotes taken from an Association of College & Research Libraries document dated 24 June 2003, found at www.ala.org/Content/NavigationMenu/ACRL/Publications/White_Papers_and_Reports/Principles_and_Strategies_for_the_Reform_of_Scholarly_Communication.htm (which makes me wonder how long it will take before a librarian-designed URL exceeds a megabyte): “Scholarly research . . . is created as a public good to facilitate inquiry and knowledge. . . . The vast majority of scholars develop and disseminate their research with no expectation of direct financial reward. . . . As journals move from print to electronic form, the legal framework for their use changes from copyright law to contract law. The latter framework governs publisher licensing agreements, which often include undesirable limits on use, eliminating forms of access that would have been permitted in the print environment under principles of fair use. . . . Mergers [of journal publishers] have resulted in documented opportunistic price increases.” The document proceeds to outline eighteen “strategies for reform in the system of scholarly communication.”

Another recent development is the Public Library of Science [www.plos.org], “a non-profit, international grass-roots organization of scientists” that has just begun this week to publish a high-profile open-access journal, *PLoS Biology* [biology.plosjournals.org], with *PLoS Medicine* to follow next year. “The full contents of every issue will immediately be placed in the National Library of Medicine’s public online archive, PubMedCentral, guaranteeing their permanent preservation and free accessibility.” The leader of this initiative, Nobel laureate Harold Varmus (former director of NIH), says “These new journals will substantially increase the value—to both the scientific community and the public—of the tremendous investment our society makes in scientific research.” Co-leader Michael Eisen of Lawrence Berkeley Lab says, “The anachronistic system of giving away the copyrights to the original research reports, and then paying for access to them, costs more, and it effectively deprives most of the world—including the people whose taxes paid for the research in the first place—from having any meaningful access to the results.” And their editorial in *PLoS Biology* goes further: “Our aim is to catalyze a revolution in scientific publishing by providing a compelling demonstration of the value and feasibility of open-access publication. If we succeed, everyone who has access to a computer and an Internet connection will be a keystroke away from our living treasury of scientific and medical knowledge. . . . The ability to search, in an instant, an entire scientific library for particular terms or concepts, for methods, data, and images—and instantly to retrieve the results—is only the beginning. Freeing the information in the scientific literature from the fixed sequence of pages and the arbitrary boundaries drawn by journals or publishers—the electronic vestiges of paper publication—opens up myriad new possibilities for navigating, integrating, ‘mining,’ annotating, and mapping connections in the high-dimensional space of scientific knowledge.”

The main idea behind PLoS is a relatively new business model for journal publishing, something like the page charges that used to be ubiquitous, under which an author’s research sponsor contributes the costs associated with open-access publication (and unsponsored research is covered by charitable donations). This model has apparently been successful in the *New Journal of Physics* [www.njp.org], published by the Institute of Physics and the German Physical Society, as well as the *Journal of Cell Science* and others. And I just found an amazing website, www.biomedcentral.com, which lists more than a hundred completely open access journals published by BioMed Central, with more coming soon; apparently they have developed some kind of turnkey system whereby groups of researchers can launch open-access journals under their own editorial control, with “no space constraints,” usually charging a flat fee of \$500 per paper. They deduct \$50 for articles formatted with certain version of bibliographic management software called EndNote. I doubt if they know how to deal with math.

The Public Knowledge Project at the University of British Columbia [pkp.ubc.ca/ojs/] has developed open-source software called Open Journal Systems, designed to “assist with every stage of the refereed publishing process, from submissions through to online publication and indexing . . . for the purpose of making open access publishing a viable option for more journals.”

John Ewing, the Executive Director and Publisher of the American Mathematical Society, has written several articles with cautionary remarks, primarily pointing out that most writers on the subject of journal publishing fail to appreciate the complexity of the situation and the relative slowness of scientists to change their habits. Particularly relevant is his note entitled “Predicting the future of scholarly publishing,” which appeared in the *Mathematical Intelligencer* for Spring, 2003, as well as in the *LNCS 2730* volume cited above; you can view it on his website, www.ams.org/ewing. “Experts and enthusiasts often dismiss past experience. They argue that because everything will soon change, experience is not relevant. . . . But they are wrong: Making predictions without facts is mysticism, not science.” He points out, for example, that *Mathematical Reviews* indexed or reviewed 51,721 journal articles in 2001, from 1172 distinct journals; 20,797 of those articles came from the non-mainstream half of those journals. In 1991, only 24% of the journals indexed by *Mathematical Reviews* were published by commercial firms, and they published 38% of the articles; by 2001, the corresponding figures had changed to 30% and 48%, indicating that journals of type (iv) are gradually coming to dominate the scholarly literature. “What drives the expansion of commercial journals? While most scholars concentrate on the *costs* of journals, *revenues* are the crucial figures in understanding journal economics. A rough estimate [by Andy Odlyzko] suggests that the revenue from each article in commercial journals is about \$4,000 (which may be off by a factor of 2). . . . Elsevier derives more than a billion dollars in revenue from its science journals.”

Indeed, the *Financial Times* in March 1997 reported that Elsevier’s profit margin in scientific publishing was 41.8%. Another article, in the *Economist*, 2001, said that the margins in their scientific and medical business were “around 35%.”

Ewing points out that the entire mathematics section of the arXiv held only about 20,000 papers in mid-2002; during the time those papers were submitted, *Mathematical Reviews* indexed more than 14 times that amount. A survey of physicists in 2001 indicated that only about 1/3 of them actually used a preprint server. It seems clear that people aren’t moving to alternative journal models as quickly as a lot of experts think. “Many scholars (although not most) promote alternatives to journals, but many fewer actually use them.” He doesn’t conclude that we should stop such newfangled experiments, quite the contrary; but he cautions that we shouldn’t expect things to change overnight. On the other hand he opines that, if it comes, “Ecological disaster for scholarly publishing would be swift and (largely) unnoticed by anyone outside academic life.” Here his expectation of swiftness seems to contradict the tenor of the rest of his article; yet a *New York Times* editorial on August 8, 2003, commenting on the Public Library of Science, was surely on the mark when it said “Most of us, admittedly, will not have much use for free access to new discoveries in, say, particle physics.” That editorial, incidentally, concluded by saying “The publishers of scientific journals are naturally skeptical, but the real test [of open access to research] will come in the marketplace of ideas. What will matter this autumn, when the new journals make their debut, is how many scientists choose to publish in them rather than in the journals traditionally deemed to be most prestigious in their discipline.”

Ewing’s website has a letter to the editor of *Nature*, dated October 2003, that expresses unhappiness with open access publishing supported by author fees. But I think the costs for mathematics are significantly lower than they are for biomedicine; and the money that libraries now pay to commercial publishers would be more than enough to cover the costs if they could be redirected. According to the Budapest Open Access Initiative FAQ, “As it is currently spent, the money buys access only for the buyer, so that access is limited to those who can afford the price. If instead, the money covers the costs of disseminating articles only, then the articles could be freely accessible to everyone. In short, the solution is to use existing funds to pay per outgoing article (dissemination), not per incoming article (access).” On the other hand, David Johnson makes the following points in a recent letter: “We in the downtrodden telecommunications industry might have trouble with the \$500 per article option. The end state in which we pay for publication rather than document delivery should [eventually turn out to] be globally cheaper, but unfortunately there would have to be a transition period in which we would need to do both. Right now we’ve already stopped paying page charges because of budgetary constraints, and in any transition period we can’t raid our continually-shrinking library budget to fund publication fees.”

I've been talking about the future, but we should also think about the past. Vast amounts of important mathematical and computer science literature now belongs to commercial copyright holders. Elsevier embarked last year on a \$40,000,000 project to scan its entire backlist—1200 titles, some going back more than 100 years. “Four sea containers and two air cargo containers packed with journals and microfiches had to be sent to the Phillipines where the scanning company is based.” The material will be archived at the National Library of the Netherlands. I commend Elsevier for taking steps to digitize all this material, but I sincerely hope that they will not abuse their monopoly on this nonrenewable resource. They own the copyright and they certainly deserve to recoup their expenses and make a decent profit; but the long-term result would clearly have been better for all mathematicians and computer scientists if we had had open-access publishing long ago. Elsevier's staff will incorporate these backfiles in the pay-per-view ScienceDirect service, which at the moment is too costly for Stanford's library with respect to full-text privileges. I've tried ScienceDirect full-text searches under Stanford's restrictions—these searches supposedly identify sources for which I could pay to get further access if I desired—but the results were quite poor. I suppose Springer will develop a similar proprietary service. JSTOR works much better at present, and tools like CiteSeer that depend on open access are significantly more useful for current research. Elsevier was a founding member of CrossRef [www.crossref.org]; if my understanding is correct, CrossRef will allow me to link to journals owned by X when I'm reading a journal owned by Y. But it won't allow me to search simultaneously through the journals owned by several different publishers. (Until all publishing firms have merged into megamath.com.) Search tools of the future will obviously be much better off if we can ultimately make the historic literature toll-free.

And now the point. Okay, I've rambled long enough. What I want you to do next is **prepare to vote** on one of four general scenarios for the future of the *Journal of Algorithms*: Should we become a journal of type (i), (ii), (iii), or (iv), according to Joan Birman's classification scheme? In the next few paragraphs I'll proceed to sketch these alternatives, in reverse order; then I'll explain how I wish to solicit your opinion via a straw ballot.

Option (iv) is to remain with our present publisher Elsevier. In this case I believe we should insist on the right for authors to put postprints on a server like the arXiv, permitting fair uses such as being crawled by CiteSeer and printed in small quantities by individual researchers; such a condition should be non-negotiable in today's research environment.

Let me try to give the main arguments in favor of option (iv), besides the obvious fact that it requires no other action on our part: First, a publisher has to do many things, and Elsevier does them well. Articles are converted from author-produced \TeX form, which often is inconsistent and unbeautiful (present company excluded), into nicely typeset pages, with expert copy editing. Authors are provided with convenient ways to submit papers via the Internet and to query their status; referees and editors are provided with convenient ways to deal with the papers. The journal is nicely printed, bound, and shipped in a timely fashion. Subscription lists are maintained; libraries are efficiently billed; non-subscribers are efficiently prevented from defacing the works. Every paper is archived, and search engines are available to subscribers. Some editorial boards have reportedly been successful at negotiating a lower price for hardcopy. Elsevier makes a hefty profit because good work deserves its reward; and the total cost of our particular journal is relatively small compared to others.

That's the "pro" side of the argument. Before I give the "con" side, let me digress a little about hardcopy. Many people seem to think that the title and abstract of a paper are sufficient to indicate whether the paper will be interesting. I strongly disbelieve that: I *always* want to flip through the pages to see what kinds of formulas and diagrams are present and to see what methods are being used. Methods are much, much more important to me than theorems. For example, consider Les Valiant's paper last year in *SIAM Journal on Computing* **31** (2002), 1229–1254. From its title, "Quantum circuits that can be simulated classically in polynomial time," and from its abstract, I could never have guessed that it has any relevance whatever to my forthcoming book on combinatorial algorithms, aka Volume 4, since that book does not deal with quantum computation in any way. But by turning pages I discovered his marvelous "matchnets," which ingeniously convert many Boolean formulas into equivalent perfect matching problems; it's a nice application to mention in my Section 7.5.5. Hardcopy, not an abstract and title, gives me the real insights that I seek. So I'm not surprised to find that, for example, the *Electronic Journal of Combinatorics* is now published also in hardcopy form as the *Journal of Combinatorics*. Some day there will no doubt be better ways to scroll through electronic documents without eyestrain, but no competition for good old paper copy is yet on the horizon.

Now to the "con" arguments re option (iv). I know that the cost of *JoA* is only a drop in the bucket, in the sense that whatever we do will not make much of a difference to the finances of our local libraries. (For example, the current cost of our journal is well under 1% of Stanford's math-and-computer-science library budget.) But it's the principle of the thing: I don't want to be responsible for any money that is being spent to keep our research proprietary. I want to keep my own house in order. Otherwise we are essentially donating content so that we and others can be charged for accessing it. If options (iii), (ii), and (i) didn't exist, and if the publisher hadn't forgotten its pledges by raising the cost of our journal, I would be all for option (iv), as I was 25 years ago. But now, I don't even see any effective way to ensure that Elsevier will keep the price of hardcopy reasonable, given that they have been unfaithful in this respect before. In the words of J. Alan Robinson, "The worlds of academic and commercial publishers seem to have grown apart to the point where incompatibilities outweigh affinities" [*Theory and Practice of Logic Programming*, Volume 1, page 1 (2001)].

Option (iii) is to switch to a nonprofit publisher, preferably a university press because of shared academic goals. Robinson wrote, "The Press in Cambridge is the oldest, and many (I am one) believe it is by far the best, university press in the world. In joining its lists the *TPLP* immediately acquires some of the patina of centuries of glorious academic excellence and devotion to scholarship. It is a splendid thing—an honor, indeed—to be associated with a publishing house which has for so long and so brilliantly served the academic community without in any way needing to compromise academic values in achieving commercial success."

In all options except (iv), we will need to change the name of the journal—Elsevier owns that—and we'll need to regain credibility so that libraries will subscribe and so that we will continue to receive an ample and sustainable flow of top-notch manuscripts. *Declaring Independence* says that "the best way to assure this is for the entire editorial board, with the name recognition of its members, to move with the journal, and for this change to be communicated broadly in your field. . . . And when you're ready to create change, SPARC can help."

Any break away from Elsevier should of course be gradual, respecting papers already in the pipeline, and allowing the *Journal of Algorithms* to come to a graceful conclusion if the publishers choose not to replace us with another editorial board. In the case of *TPLP* some authors did, however, withdraw manuscripts that they had submitted to *JLP* so that they could resubmit them to the new journal.

The main argument I can think of against option (iii) is that nonprofit publishers still charge rather more than options (ii) and (i). Perhaps they do this to subsidize other worthwhile activities; but such arguments don't sway me. Do they produce a substantially better product? And are they open-access?

Krzysztof Apt told me in a recent email that “the Cambridge University Press journals are not open access. However, in the case of *TPLP* we had no problem securing their agreement that the final versions of the papers are posted at arXiv and the authors are strongly encouraged to do so at the moment their paper is accepted for publication.” He also remarked that, in 1999, the publisher of *JLP* (Elsevier) did not agree to any such posting.

Option (ii), in our case, would almost surely be to link up with ACM and to morph into *ACM Transactions on Algorithms*. We would make a proposal to ACM’s Publications Board, stating which people have agreed to serve as editors and how our mission relates to *JACM* and other ACM publications. I’ve heard of various informal proposals that indicate the strong viability of such an option. ACM does its transaction volumes just as well as Elsevier does its journals, with respect to typesetting, copy editing, paper shuffling, subscriptions, archiving, timeliness, etc.; and with option (ii) I believe many more people would acquire paper copies of our journal in their own homes or offices than with the other options, because ACM sends out convenient order forms when people renew their SIGACT membership. For example, I subscribe to *TOPLAS* but not to *Algorithmica*.

An alternative for option (ii) would be to head towards a *SIAM Journal on Algorithms*. In my opinion SIAM doesn’t have its act together as well as ACM does at the present time, especially with respect to copy editing and typesetting. (They haven’t even upgraded the fonts that I’ve been complaining about on my website for years at www-cs-faculty.stanford.edu/~knuth/cm.html; I’ve given up trying to nag them about this, it’s evidently hopeless.) SIAM’s pricing structure is also flaky, as noted above.

I wish it were true that ACM transactions volumes qualify as “open access,” because the Computing Research Repository page [xxx.lanl.gov/archive/cs/intro.html] says that CoRR is sponsored by ACM and the arXiv.org e-Print archive and others. But clicking “Find out more about CoRR” on that page gets me no information whatever. The current ACM copyright policy, adopted in November 2002, can be found at www.acm.org/pubs/copyright_policy; it is no more permissive than Elsevier’s, and significantly less so than that of Cambridge University Press: “2.5. . . . The right to post author-prepared versions of the work covered by ACM copyright in a personal collection on their own Home Page and on a publicly accessible server of their employer . . . is limited to noncommercial access and personal use by others, and must include this notice both embedded within the full text file and in the accompanying citation display as well: ‘© ACM, YYYY. This is the author’s version of the work. It is posted here by permission of ACM for your personal use. Not for redistribution. The definitive version was published in . . .’ . . . Authors may post works on public repositories before acceptance but must incorporate the ACM copyright notice upon transfer of copyright. After acceptance, authors may post the work on public repositories only with the explicit permission of ACM.”

Option (i), in which we enlist a university or group of universities and/or research institutes to host our journal, is the most radical change. But it is the option most clearly in the open-access camp, and there now are many existence proofs of its viability and its cost-effectiveness for everybody. If we’re going to change, perhaps we ought to jump right into the future instead of taking baby steps. With this option I suspect we would find it easiest to take advantage of new opportunities offered by the Internet, like the use of color and sound, and the introduction of applets into interactive papers, etc.

Jim Pitman at Berkeley has been thinking seriously about such issues with respect to statistics journals, and he just sent me the following email message: “I think the model we should aim for with our journals is this: Set up the production to be efficient enough that the organization producing the journal can afford to post electronic copies of final versions of papers, for free online access via arXiv or similar open access digital repository. Then find a paper publisher who is willing to print and distribute paper copies of the journals on this basis. I expect that for some time there will be enough demand for paper copies that libraries will continue their subscriptions even though the content is free online. If you are willing to sacrifice typographical quality just a little, i.e., let authors provide final $\text{T}_{\text{E}}\text{X}$, there is no question you can support electronic production on zero cash flow. EJP/ECP proves this. [He means the *Electronic Journal of Probability* and *Electronic Communications in Probability*, hosted by the University of Washington math department: see www.emis.de/journals/EJP-ECP/submit.html.]

G&T is a good model with higher quality typography and a low cost paper version. I plan to follow their model for the Institute of Math Statistics journals.”

I personally would not be associated with a journal that does not provide copy editing and typesetting services; it pains me to read raw conference proceedings. But I’ve heard that Elsevier and Springer farm these activities out to entrepreneurs in India and China who do an excellent job at a reasonable price.

How to vote. As I said, the purpose of this letter is to solicit your opinions, hence I’m interested in the results of a “straw ballot”—which is for information only, not bound to any definite action.

Namely, I’d like each of you to rate the desirability of each of options (i), (ii), (iii), and (iv) as described above, in the following way: You have 10 votes to be distributed among the four categories, representing your relative preferences. For example, you could give four votes to your favorite option, three to the next, . . . , and one to your least favorite; or you could put all ten votes on a single choice. Whatever.

But *do not vote immediately!* Please take some time to think about the issues, to look at samples on the web, to talk things over with colleagues. I think Zvi Galil and David Johnson will be writing to you shortly, giving their own insights; I’ve already shown this letter to them, but I had decided in advance that I wanted to express my opinions in a personal and informal letter, for which I could take sole responsibility for mistakes and biases, because I feel so deeply about our journal. I have also sent a copy of this letter to our representatives at Elsevier, because they might well want to present their side of the story better than I have done.

I have listed lots of facts above, and I’ve probably not gotten them all straight. I’ll write you again to correct any errors that are brought to my attention.

Although I don’t want you to rush, I do ask that you get your vote to me by Monday, November 24. I usually shun email, but in this case I won’t mind your message addressed to `knuth-bug@cs.stanford.edu`. Please cast your vote *before* knowing how anybody else will respond, so that the votes will be independent. I suggest *not* emailing comments back and forth between the board members as a group, until the straw vote has been taken.

After receiving everybody’s ballots, I’ll report the results back to you, not identifying anybody by name but rather stating how many people voted in each of the various ways. Of course I’ll also compile any further comments that you might wish to share with the group. Then, after all votes are in, we can decide what the next step ought to be.

This has been a terrifically long letter, and I’ve quoted many thoughts from people who have pondered the questions more deeply than I. Let me close with one more, taken from Harold Boas’s editorial in the *AMS Notices*, September 2001, page 789: “Scientific researchers dedicate their lives to building up the edifice of human knowledge. . . . Once publishers recoup the expenses they incur in helping to paint the structure, they should then relinquish any claim to ownership of the building. . . . ‘Give me a place to stand,’ Archimedes is supposed to have said, describing the principle of the lever, ‘and I will move the world.’ How much leverage will we apply to move the world toward online public repositories of our common scientific knowledge?”

Sincerely,

Don Knuth

DEK/dek

P. S. I’m sending copies of this letter to several friends who are interested in journal publishing but are not members of our board. But this is not an “open letter”; I would prefer *not* to have my remarks circulated widely. I’m emphatically not a revolutionary. I just want our journal to do the right thing.