Top Ten Reasons to *Not* Share Your Code (and why you should anyway)

Excerpt prepared for the ICERM Workshop on
Reproducibility in Computational and Experimental Mathematics
http://icerm.brown.edu/tw12-5-rcem
December 10-14, 2012

Randall J. LeVeque[1]

This thought piece is really a request to read the draft of an article I am working on for *SIAM News* and give me some feedback.

The article can be found at http://faculty.washington.edu/rjl/pubs/topten/.

This article grew out of a talk I gave with the same title in a minisymposium at the 2011 SIAM CSE meeting in Reno, organized by Jarrod Millman, on "Verifiable, Reproducible Research and Computational Science". (Slides from my talk and others are available at http://jarrodmillman.com/events/siam2011.html.)

The gist of the article is to urge readers to reconsider current attitudes about sharing code related to publications by pondering an "alternative universe" in which mathematics papers are not expected to contain the proofs of theorems. Many of the objections I hear repeatedly to sharing code can be applied to such a universe.

As a teaser, here I will just list the reasons people might object to sharing proofs of theorems in this alternative universe. Obviously sharing code has some issues that differ from publishing proofs. Before you point these out to me, please read the full article where I attempt to deal with some of them.

One thing I do want to point out, which I stress in the article, is that to me sharing code to accompany research papers in applied mathematics is more about being able to inspect the relevant part of the code than about being able to run the code and produce exactly the same answers. Parameter choices or implementation details that may crucial information for the reader or referee is frequently missing from the article. Personally I think people get hung up too much on the fact that it's hard to insure others can run the code, and should focus more on providing a full record of the research methodology.

[1] Department of Applied Mathematics, University of Washington, Box 352420, Seattle, WA 98195-2420, rjl@uw.edu.

Top ten reasons not to share your proof:

1. *The proof is too ugly to show anyone else.* It would be too much work to rewrite it neatly so others could read it. And anyway it's just a one-off proof for this particular theorem, and not intended for others to see, or to use the ideas for proving other theorems. My time is much better spent proving another result and publishing more papers rather than putting more effort into this theorem, which I've already proved.

2. *I didn't work out all the details.* Some tricky cases I didn't want to deal with, but the proof works fine for most cases, such as the ones I used in the examples in the paper. (Well, actually I discovered that some cases don't work, but they will probably never arise in practice.)

3. *I didn't actually prove the theorem, my student did.* And the student has since graduated, moved to Wall Street, and thrown away the proof, since of course dissertations also need not include proofs. But the student was very good, so I'm sure it was correct.

4. *Giving the proof to my competitors would be unfair to me.* It took years to prove this theorem, and the same idea can be used to prove other theorems. I should be able to publish at least 5 more papers before sharing the proof. If I share it now my competitors can use the ideas in it without having to do any work, and perhaps without even giving me credit since they won't have to reveal their proof technique in their papers.

5. *The proof is valuable intellectual property.* The ideas in this proof are so great that I might be able to commercialize them some day, so I'd be crazy to give them away.

6. *Including proofs would make math papers much longer.* Journals wouldn't want to publish them and who would want to read them?

7. *Referees would never agree to check proofs.* It would be too hard to determine correctness of long proofs and finding referees would become impossible. It's already hard to find enough good referees and get them to submit reviews in finite time. Requiring them to certify the correctness of proofs would bring the whole mathematical publishing business crashing down.

8. *The proof uses sophisticated mathematical machinery that most readers/referees don't know.* Their hardware/software cannot fully execute the proof, so what's the point in making it available to them?

9. *My proof invokes other theorems with unpublished (proprietary) proofs.* So it won't help to publish my proof — readers still will not be able to fully verify correctness.

10. *Readers who have access to my proof will want user support.* Anyone who can't figure out all the details will send email requesting that I help them understand it, and asking how to modify the proof to prove their own theorem. I don't have time or staff to provide such support.