ACL 2017

PC CHAIRS BLOG

Last minute reviewing advice

knmnyn *Uncategorized* *6 Comments*

Dear reviewers (and readers),

The reviewing deadline for ACL submissions is a few days away. In line with our previous post on *Last minute writing advice (http://chairs-blog.acl2017.org/2017/02/02/last-minute-writing-advice/)* that featured senior NLP/CL personalities, we've solicited more advice on the path to rewarding and meaningful reviewing from another cadre of excellent reviewers*, on how to write useful reviews and their pet peeves about poor reviews that they have received in the past.

As always, we hope this post will pique your interest in contributing! Since you're reviewing right now, if you have more thoughts to add to the criteria here, for posterity, please comment below!

– Regina and Min

* Most of our post contents come from the well-qualified role of outstanding reviewers from the reviewers' role in the NAACL 2015 frontmatter (http://www.aclweb.org/anthology/N/N15/N15-1000.pdf).

Mirella Lapata (http://homepages.inf.ed.ac.uk/mlap/) (University of Edinburgh)

What makes a great review? Well, it's easier to explain what makes a bad review.

1) *A review based on a very superficial reading of the paper which asks questions addressed possibly even in the introduction*. These reviews often tend to be short (e.g., "I really liked this paper, it addresses and important and timely problem") and vacuous.

2) *A review based on sentiment and ideology rather than on the merits of the paper*. For instance, I do not like LSTMs and thus dislike every paper using them (obviously this is just an example, who doesn't like LSTMs?)

3) *A review which claims things about the paper which are unsupported by evidence*. Statements like "this has been done previously and is therefore not novel". Who did it, when, and how?

4) *A review which says great things about a paper but offers low scores*. If the paper is no good, then the discourse of the review must correspond to the numeric evaluation. Be constructive rather than polite.

5) *A lukewarm review*. This is the majority of the reviews in *CL conferences, the vast grey area of papers represented by the number 3. I don't have a solution for this, maybe get rid of 3 as a number 🙂 But seriously, reviews that take a stance are more useful to everyone involved.

Finally, reviewers out there, bear in mind authors are human beings, they have feelings too! The point of the review is not to annihilate but rather support and improve scientific communication and the dissemination of research.

Marco Baroni (https://research.fb.com/people/baroni-marco/) (Facebook Artificial Intelligence Research)

When reviewing, I try to apply a piece of advice I received from Bruce Hayes, my thesis supervisor: Avoid vague language. That is, you should not use expressions such as: many, much, several, "for example", "and others". Consider the statement: "The idea in this paper is a minor development of what was proposed by Pinko (2001) and others". This statement is reinforcing the claim that the paper is not original by referring to some unspecified "others", which is doubly unfair, as the authors cannot defend themselves by explaining how their work is different from that of these mysterious "others". As another example: "The exposition in the paper is too informal. For example, the way in which the input images are embedded in the output space should be made explicit with an equation". Again, this claim suggests that there are many other ways in which the paper is too informal, but it doesn't say which. I find that, by applying the avoid-vague-language rule, I often end up having a fairer/more positive opinion of the paper I am reviewing (for example, I might realize that my initial impression that the paper is confusing is actually due to just one or two paragraphs that are effectively confusing).

Yoav Artzi (http://yoavartzi.com/) (Cornell University)

There are many causes for poor reviews. In conferences, one cause is narrowing the process to only writing the review. Treating reviewing as a dialog is key to clarify reviewer confusion and refine the criticism. This is a dialog between the reviewers, but also the authors. I find discussion between reviewers useful to understanding muddled parts paper, and better identifying its strengths and weaknesses. My opinion of papers has gone both ways following discussions. Not less important is the author response, which can drastically increase the review quality. Critical to this process are: (a) reviews that list core concerns clearly and present questions to the authors, (b) authors that address the concerns in their response (instead of getting defensive and taking it personally), and (c) reviewers that go back to the response, read it carefully, and, if necessary,

modify their scores. I hope that this year's longer discussion period will drastically improve reviewing quality. Even when papers are not accepted, good reviewing make their next version significantly better.

What do I like to see in a great review? It's always re-assuring to see the reviewers understood. A one-sentence summary, although redundant, is a positive start. The core of the review should be the major issues with the paper, if there are any (there usually are). These should be clear and to the point. A really great review will also detail **separately** minor flaws and comments that will just make the accepted paper better. A minor tip I try to follow: focus on the paper, not the authors. Addressing the "paper" and the "approach" rather then "the authors" is maybe just cosmetic but goes a long way (I believe I picked it from Hal's blog (http://nlpers.blogspot.sg/2014/04/an-easy-way-to-write-less-hurtful.html) at some point).

Emily M. Bender (http://faculty.washington.edu/ebender/) (University of Washington)

There are only two audiences (that matter) for your review:

- The program committee (area chairs, PC co-chairs)
- The authors of the paper

Keep these audiences in mind as you write the review. What you write should be helpful for both of them. Specifically:

(1) Start your review with a summary of the main contribution of the paper. This helps the ACs because it contextualizes the rest of your comments and it helps the authors because it lets them know whether you have understood the paper. It can also be helpful for authors to see how someone else (with likely at least a somewhat different perspective on the field) phrases the main contribution.

(2) *Be Specific*. Vague general statements, positive or negative, don't help the ACs make an informed choice about which papers to include in the program and also don't help the authors improve their work. If you think the paper is missing critical references, give the actual citations. If you think the evaluation or error analysis is insufficient, state specifically what is missing/what needs to be changed.

(3) *Be Kind.* Keep in mind that there's a person (or people) on the other side of the double-blind review process, with feelings. That person might be a student or someone else new to our field whose contributions we will benefit from in the future, so there is no gain in writing a nasty review that scares them off. Negative reviews can still be phrased kindly. This can be done by taking the time to find something to praise even while recommending that a paper be rejected (e.g. "The question this paper asks is an important one. However, the methodology described doesn't convincingly answer the question because…") or by rereading and editing for tone, asking yourself "How would I feel if I or my students received this kind of review?" Similarly, if you see co-reviewers being unduly mean, point it out ask them to rephrase.

(4) *Be Honest*. Describe your own reactions to the paper. Statements like "I had a hard time following Section 3", or "I was confused about X until I got to 4.2; it would help to have it explained briefly in Sec 2" are more helpful to authors statements that purport to be objective like "The paper is confusing." It is also helpful to accurately represent your own expertise: "I'm an active researcher

in this field, and yet I still found Sec 3 confusing" or "This is not my area, but I would still expect to be able to understand Sec 3 better than I did." Similarly, honest questions can be a good way to convey what is missing in the exposition: "What metric do the numbers in Table 1 represent?"

(5) *Be Optimistic*. Reviewing that looks for reasons not to accept papers leads to conferences full of incremental work. Make a concerted effort to articulate what you like about the paper, in addition to how you think it could be improved. Outstanding literature review? Say so. Novel approach to error analysis? Likewise. Clever application of existing technique in a new domain? Show that you noticed. Such feedback helps authors not just with morale but also by positively reinforcing good research and good writing. It also helps the program committee construct a varied and worthwhile conference program.

Joel Tetreault (Grammarly)

My general piece of advice when writing a review is to think about who is going to be reading the reviews and what they hope to get out of it. In this case, the recipients are the area chairs, other reviewers and authors – in no particular order.

For the area chairs and the other reviewers, a few things are important: be organized, be clear, and be as decisive as you can. Put yourself in the shoes of an area chair: each area has anywhere from a dozen to well over a hundred papers submitted, which can mean that an area chair has potentially hundreds of reviews to sift through in a short period of time. So, having a review that is clear and cohesive is very important to making their lives easier and also making well-informed decisions. What can help is to organize your review into pros and cons sections and then explain why you assigned your scores a certain way. For instance, a paper might not offer anything new in the way of algorithms (thus undercutting novelty) but it might introduce a new task and dataset which you may view as important. You can advocate for acceptance stating that the strengths of the latter points outweigh deficiencies in some aspects of novelty. Obviously there will be cases where it can be hard to make a decision, but you can still make explicit why this is a hard decision. Another way of organizing your review in a helpful manner is to place your main arguments for and/or against the paper in the beginning of the review and then minor comments and questions at the end. The above comments are also related to assisting in the reviewer discussion period. If the other reviewers can see your main points and where you are coming from more easily, it makes for a more enriching and even enjoyable discussion session.

For the authors, a well-organized and coherent review makes their lives easier in terms of understanding what motivates your recommendation as well as addressing your comments in the author response. I want to add three more points to keep in mind. First, be nice. Our field has grown rapidly in the last few years and we have many junior researchers entering the fray, researchers from other fields publishing in the ACL for the first time, and non-native speakers of English who may write beautifully in their native language but may not have yet reached the level of a native speaker. A community is known by how it treats its members, and reviews are a part of that. So even if a paper has significant issues, make sure to communicate criticisms in an even-handed manner. Second, be constructive. Instead of simply enumerating all the criticisms and drawbacks to the paper, try to frame your feedback by answering the question "what can I, as a reviewer, do to assist the authors in improving the paper?" Third, be detailed. For example, if a paper has missed some key related work, make sure to provide the citation or even the URL. One of the best reviews I ever read as an area chair was from a very senior professor who didn't advocate

for acceptance but still listed several great ideas in their review to help the authors out. The outcome may have been negative for the paper at that time, but the review was encouraging, and the authors came out of it with next steps for their work.And finally, the most important point to remember: don't leave your reviews to the last minute. 😉

Ani Nenkova (University of Pennsylvania; NAACL 2016 co-chair)

Last year Owen Rambow and I wrote a note to the area chairs just before discussion period. By understanding the whole process better, hopefully reviewers can write better reviews.

In the post here I would like to talk about reviewing work in emerging areas of computational linguistics. Stuff we have seen before sounds more convincing and is easier to evaluate. New topics, tasks and questions are harder to evaluate and could be adopted with much uncritical enthusiasm or readily dismissed as incomplete, unimportant or imperfect. Both are bad for the field.

If a new task with flawed definition or a problematic data makes in into the conference, others may follow up on the task and waste much time, effort and other resources. If papers presenting new directions for research get rejected, the field risks loosing those authors and losing an opportunity to evolve in a natural way, rather than borrow topics and ideas that have already appeared in related conferences on machine learning, vision or AI. The reasons for rejection of such unusual papers is often "it's not important" (maybe the reviewer doesn't see the importance now), "it is not NLP" (but may be NLP in five years), "there are many questions left unanswered" (which indicates that there is much room for follow up work by the authors and others).

I have no solution on how to make the best judgement but I hope that reviewers will take extra time and thought when making a recommendation on a paper on emerging or new topic. If the paper asks a clear question, has well-designed experiments to answer aspects of this question and draws conclusions consistent with the experimental results, the paper most probably belongs in the conference program.

Tim Baldwin (University of Melbourne)

• *give as good as you would like to get*: it's great when you get a good, insightful review that greatly improves a paper; try to be that reviewer with your own reviews, and keep an internally high standard with your own reviews, and raise the overall bar in the process.

• *if the glass was half (or more) empty, what would it have taken to make it full?*: pointing out the deficiencies of a paper is an inevitable part of reviewing, but a good review equally provides suggestions of what you feel could/should have been in the paper to get it to a level where you would have been comfortable accepting it.

• *secondary reviewing should be a great opportunity to mentor, not a simple exercise in outsourcing*: for mentors/supervisors of students, secondary reviewing can be a great way of breaking your mentees into reviewing, but make sure to go over the paper with them, form your own impressions, provide feedback on a draft review from them, and get the final review to a point where you are prepared to defend it in the discussion phase; that is, by all means make use of secondary reviewers, but expect to have to put as much or more work into it than if you reviewed the paper yourself.

• *accept that the distribution of papers that you review won't necessarily be representative of the true population*: a common trap to fall into is the mentality of assuming that one of your assigned papers will be worthy of acceptance but the other 3 (or whatever number) won't, i.e. that the distribution of your reviews will mirror the overall acceptance rate; it may, of course, be possible that you get assigned all great papers or all awful papers, and the important thing is to assess each paper independently on its own merits.

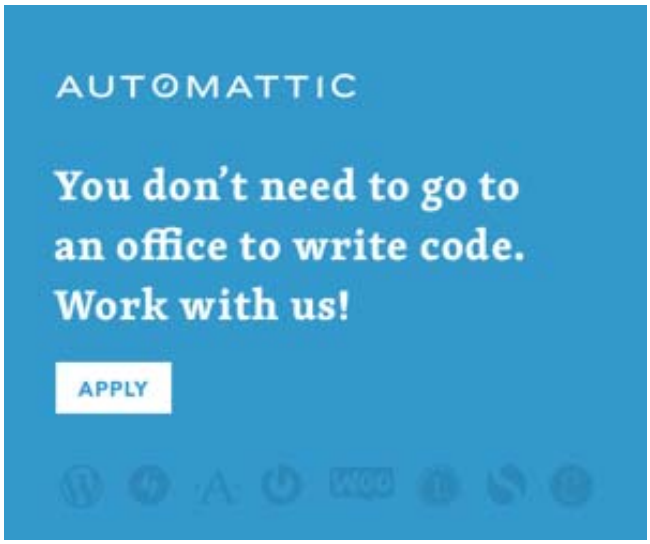
• *allow time to digest the paper, and let your review settle*: sometimes, you will only appreciate the true worth of a paper/pick up on flaws in the paper after a bit of digestion, and as such, you want to avoid doing your reviews at the very last minute, to allow time for this to occur; equally for your reviews, you ideally want to be able to leave enough time to read back over your review to make sure it is clear, and potentially smooth over any inadvertent negativity/snark (there's always a constructive way to frame any criticism).

• *reviewing is not (primarily or even secondarily) a beauty contest: yes, a well-formatted paper with excellent presentation is an easier and more pleasant read, but try to see past the presentation of a badly-formatted/awkwardly-presented paper to see if it has technical/scientific merit, and equally, don't allow yourself to be dazzled by a nice-looking, easy-to-read paper, or be unduly biased in your scientific assessment of the paper.*

• *avoid unsubstantiated blanket statements: we've all been on the receiving end of lacklustre reviews which state that our paper is "poorly written" or "underdone" without any further substantiation; if this is your assessment of the paper, make sure to justify the comment with examples of places where the paper is poorly written (and how you would have liked it to have been written), or specifically in what way you feel it is underdone (and what you would have liked to see to be convinced by the paper — see above).*

• *engage with your fellow reviewers and the paper author(s): the author response and discussion phases serve an important purpose, and can provide invaluable extra insights into a paper to the Area Chairs and Programme Chairs; this is an integral part of your reviewing responsibilities, and it is critical that reviewers engage in this process.*

• *hold onto those jottings for the discussion phase: the lag between submitting your reviews and engaging in the discussion phase is often long enough that it is hard to remember the subtler points in your own review; having your original jottings can be very helpful in recalling the details of the paper and reminding yourself of your arguments for/against a paper.*



6 thoughts on "Last minute reviewing advice"

1. **RON ARTSTEIN**

   says:
   February 24, 2017 at 5:41 am
   Thanks everyone for all the great points (and apologies to Regina and Min for not getting my ideas in time). A few additions:

   – When asking for changes in accepted papers, keep in mind constraints of time and space. Providing some extra data is often feasible within the timeframe for revisions; running an intensive computation or a new user study is usually not. ACL is good about giving authors extra space for addressing reviewer comments, but even this space is limited; if a large amount of extra detail is needed, it's helpful to also recommend places in the paper that can be trimmed down.

– Remember that the difference between long and short papers is one of scope, not quality. Don't recommend shortening a paper unless the research is of high quality and also of small enough scope that it can be adequately expressed in a short paper. (One of the most frustrating things a reviewer can do is simultaneously ask for more detail and recommend shortening a paper.)

And a question for Regina and Min: Traditionally ACL has separate deadlines for long and short papers, so the question of changing a paper's category doesn't arise. At some conferences which review long and short papers simultaneously, requests for changing a paper's category are routine (usually from long to short). So what is the policy for ACL this year? Are reviewers allowed to recommend changing a paper's category? Should the area chairs or program chairs entertain such requests?

1. **SADID A. HASAN**

   says:
   February 28, 2017 at 2:07 pm
   Great point Ron. While reviewing I also felt that some long papers don't have enough contributions, but may suit well as a short paper.

2. **KNMNYN**

   says:
   March 2, 2017 at 8:30 am
   Hi Ron, thanks for your question and sorry for the delay in the reply. No, we are not supporting changing from long to short (or vice versa). We want authors to take a stand on whether they also feel that their work merits a long or short format.

   *In many ways, it is more difficult to properly compose and frame a short paper. We believe short papers are better as well thought out instances of short papers, rather than crippled long papers. We'd prefer that authors whose long papers are felt by reviewers as more suited for the short paper format, to retool their papers for a subsequent *ACL event (EMNLP is coming up). This gives the reviewers a chance to vet and approve such changes, rather than relying on the authors to make the substantial revisions that would be needed to make a great short paper.*

2. **MAJA**

   says:
   February 24, 2017 at 12:56 pm
   Very good advices, thanks to everyone!

   As for author's response, I find it very useful that this year there is the option "can author response influence your scores and how?".

   As for the score "3", I support banning 🙂 (or using scores 1-6, in that way three are low and three are high, no real middle ground).

   1. **KNMNYN**

      says:

March 2, 2017 at 8:36 am
Hi Maja, I agree partially with your comment (and Mirella's) that a '3' is a unhelpful response.

I differ on a few points.

Sometimes '3' belies the implicit 'standard deviation' that would come with it. Is it really a lukewarm contribution, incremental advance? Or possibly a contentious issue where some critical part is not properly grounded?

I'd say we'd like to encourage reviewers to tend towards extremes to help paint a positive or negative signal in accepting and rejecting submission, but '3' is useful at the initial stage, when reviewers and authors are working together towards an understanding of the submission. Hopefully with the discussion period, fellow reviewers can sharpen their critiques in a way that is helpful for each other to decide which way they (individually) want to lean.

2.

C. GÓMEZ-RODRÍGUEZ (@CARLOSGR_NLP)

says:
March 3, 2017 at 10:57 am
I don't see the point in banning the score 3. There are papers that are genuinely on the fringe. And as reviewers we don't know what other papers are competing with them, so the most accurate thing to say might be "it's on the fence, it depends on what other papers are there". Nothing wrong with that.

If the 3 were banned, then many reviewers would be torn between 2 and 4 on many papers (I know I would) and the decision would depend on factors like what other papers the reviewer got, or just plain old good/bad humor. Scores and acceptance/rejection would be much more random.

"Everything Should Be Made as Simple as Possible, But Not Simpler" – to me, banning the 3 would be an example of making it simpler.

COMMENTS ARE CLOSED.