

# Notes from EMNLP-IJCNLP 2019

## Program Co-Chairs

### Major Challenges

- Recruitment of reviewers: Although we anticipated 3000 submissions and asked SACs to prepare enough names in the candidate list, we didn't expect that nearly half of the reviewer invitations were declined.
- Recruitment of SACs/ACs: This also took much longer than we expected, so it's good to start it early.
- Paper assignment: Because of the shortage of reviewers, we had to do a lot of manual work before we told SACs to assign papers to reviewers. We had to load balancing by shifting papers out of areas lacking reviewers to areas with extra reviewers, and we had to ask some reviewers whether they could split their load across two areas. These tasks took us a lot of time.
- Softconf system: The system was not designed to handle the large number of submissions we are facing nowadays. Watch out for the following: (1) Submission deadline: Make sure to set the actual deadline to be at least 3 hours after the official deadline, and make sure to configure the settings early, at least a few hours before the official deadline. Once the system encounters heavy traffic and is down, even program chairs cannot log in to fix it. (2) Sending emails: The system often crashes when we need to send hundreds of emails. Our workaround was to send important emails by tracks (which required us to repeat the same procedure  $18 \times 2 = 36$  times!) and to also bcc the emails to ourselves so we could make sure all important emails were sent.
- Handling emails: We received many emails from authors and reviewers related to different issues. E.g., requests for review load reduction, requests to add/remove authors after acceptance, etc.

### Ethical Issues

This problem didn't occur to us until some reviewers expressed concerns. It would be good to have clear guidelines from SIGDAT for future conferences. Guidelines should be given to authors, area chairs and reviewers.

# Timeline

Item	Date	Comment
SAC invitation	Feb 6 - Feb 28	It took almost a month.
AC nomination by SACs	Mar 1 - Mar 22	
Area chair invitation	Mar 23 - end of Apr	
Reviewer invitation	First round of invitations were sent at the end of Apr. The process continued even after the submission deadline.	We intended to invite reviewers earlier, but due to the delay in SAC/AC recruitment, we couldn't do it earlier.
Abstract due (Long & Short)	May 15	We wanted to use this to help us estimate the number of submissions in each area so we could recruit more reviewers or shift reviewers during the one week window, but it didn't turn out to be very useful. One reason is that many abstract submissions didn't turn into full paper submissions.
Submissions due (Long & Short)	May 21	Make sure to leave the system open for at least a few more hours beyond the official deadline.
Desk rejects, resolving COIs, <b>recruiting more reviewers, shifting reviewers/papers</b> , etc.	May 22 - Jun 4	
Bidding period	Jun 5 - Jun 8	We had to give very short bidding period because bidding was seriously delayed. And this was mostly because of the shortage of reviewers.

Paper assignment by SACs	Jun 8 - Jun 16	
Review period	Jun 17 - Jul 5	Review period had to be shortened, again because of the delay earlier.
SACs/ACs to chase late reviews, look for emergency reviewers, etc.	Jul 5 - Jul 9	
Author response period	Jul 10 - Jul 15	<p>We had to continue to look for emergency reviewers (again it's not easy because of the shortage of reviewers), and for some papers we extended their author response period.</p> <p>However, in retrospect, having an author response period gave authors certain expectations from the reviewers and ACs, which are hard to meet.</p>
Discussion and meta reviews/recommendations from ACs	Jul 16 - around Jul 26	
SACs to make recommendations	Jul 26 - around Aug 5	
PC deliberation	Aug 6 - Aug 11	
Notification of acceptance	Aug 12	It might be good to ask authors NOT to check the START system and wait for the emails.

		Alternatively, instead of sending emails, make final reviews and decisions visible to authors in the system.
Camera-ready due	Aug 30	Many authors would update their camera-ready after the camera-ready deadline. It depends on how flexible the publication chairs are. Our publications chairs were very kind to take in some last minute updates.
Main conference	Nov 5 - Nov 7	

## Softconf/START

The conference system was not robust when receiving thousands of submissions and sending emails to thousands of authors. We had to extend the submission deadline by 3 hours (and this may still not be enough next year). When sending notifications, we had to send emails to the authors in each track separately, which was very time-consuming. The system needs to be tweaked for specific things we needed, and we encountered quite a few non-trivial problems when using the system. We had to contact Rich from softconf many times. Good thing is he usually responded very quickly. An interesting tool implemented in the system is the DUDE Plagiarism Check tool, which can perform a plagiarism check against a set of papers in a database. We have used the tool to detect duplicate or significantly overlapping submissions to EMNLP or dual submissions to EMNLP and other conferences like NeurIPS.

Overall, the suggestion is to explore the functionalities of Softconf early, never hesitate to contact Rich from softconf, but to also double check the functionalities after Rich's changes.

## Program Committee

We adopted the hierarchical structure of PCs, Senior Area Chairs (SACs) and Area Chairs (ACs) to handle the large volume of submissions while maintaining a high-quality review process.

## SACs/ACs

Similar to EMNLP 2018, in this year's conference, the SACs basically acted as program chairs within their tracks; we delegated to them nearly all communication with and coordination of ACs. We recruited only one SAC for each track and thus reduced the communication overhead. For each track, both PCs and SAC need to recommend AC candidates because we need many ACs to handle the large number of submissions. Finally, we had 18 SACs and 152 ACs in total. We never had any live meeting with either the SACs or ACs.

## Reviewers

We centralized the reviewer invitation process. We asked the SACs to collaborate with the ACs and come up with the reviewer list for each track (but we also provided reviewer lists from recent NLP conferences, which we collected from the PCs of those conferences). This turned out to be a very slow process (similar to EMNLP 2018).

We used START's system for inviting reviewers for each track. We sent out the invitations from different areas at the same time to avoid different areas competing for reviewers. A reviewer may receive multiple invitations from different areas but we strongly encouraged them to choose only one.

The reviewing load for each reviewer was no more than 6 in total, and if a reviewer accepted the invitation from multiple areas, we assumed that he/she were willing to take the full reviewing load of 6 papers in EACH area. (We noted this in the invitation email and some reviewers were indeed willing to review more than 10 papers). However, this also created misunderstanding (as people didn't read our invitation email carefully) and some reviewers complained about having more than 6 papers. So in the end, we still assumed a total load of 6 for reviewers who joined multiple tracks. By default, we didn't ask each reviewer about how many papers he/she wanted to review. But there were still many reviewers requesting a reduced load. As we couldn't afford to lose these reviewers altogether, we agreed with their requests, but this also made paper assignment more difficult.

Some tracks had enough reviewers and some other tracks had insufficient reviewers, and thus we need to do review load balancing across tracks, which was very time-consuming. Many reviewers reviewed for two or three tracks and we need to split their total review load across multiple tracks.

## COI Papers

PCs were not allowed to submit papers to the conference. SACs and ACs were allowed to submit their papers to the same tracks they oversaw. We marked all the SAC COI papers and

the AC COI papers in the system so that the SAC/AC cannot see their own papers. For SAC COI papers, we tried to move them to other relevant tracks. A few SAC COI papers were still kept in the original track if we cannot find another suitable track, but they were overseen by PCs. We (PCs) assigned the SAC COI papers to ACs and informed the ACs about the SAC COI papers they handled. ACs can only discuss these papers with us. For AC COI papers, each of them was handled by another AC in the same track. Unlike EMNLP 2018, we didn't move all COI papers to a different track because it is hard to find suitable reviewers from a different track for many COI papers. (The tracks are generally broad and thus more independent with each other). Also, some SACs expressed concerns with their papers moved to other areas. The COI marking function in the system is very useful and we set COI papers for SACs, ACs and reviewers both automatically and manually.

## Submissions

EMNLP-IJCNLP 2019 adopted ACL's new policies for submission, review, and citation. Submissions which violated any of these policies were rejected without review. Note this does not apply to demo papers, which were subject to single-blind review.

Most importantly, the policies define an anonymity period, which started on April 21, 2019 (11:59pm PDT).

## Multiple Submission Policy

We made a major change regarding this year's multiple submission policy. EMNLP-IJCNLP 2019 did not consider any paper that was under review in a journal or another conference at the time of submission. This policy covered all refereed and archival conferences and workshops (including ACL workshops). For example, a paper under review at an ACL workshop cannot be dual-submitted to EMNLP-IJCNLP 2019. In addition, we did not consider any paper that overlapped significantly in content or results with papers that would be (or had been) published elsewhere. Papers may not be submitted elsewhere during the EMNLP-IJCNLP 2019 review period. Authors submitting more than one paper to EMNLP-IJCNLP 2019 must ensure that the submissions do not overlap significantly (>25%) with each other in content or results.

Due to a significant increase in submissions to recent NLP conferences, we had to change the policy in order to reduce the workload for reviewers. As we know, this was the first attempt to forbid dual submissions in ACL/EMNLP/NAACL conferences, although conferences in related communities (e.g., SIGIR, ICML, NeurIPS, etc.) have been banning dual submissions for many years. Some authors, especially those who are used to submitting their papers to EMNLP and another conference (e.g., INLG) concurrently, had complaints about this new policy.

## Instructions for Authors

We decided that the author list could not be changed after the paper submission deadline. We believe that this could make the reviewing process fairer because all COIs could be detected. There were still many requests for adding/removing authors or changing author order after the paper submission deadline, especially after the paper acceptance notification. For most such requests, we said no, while we agreed to change author order.

## Abstract Submission

Authors were asked to submit abstracts one week in advance. We initially wanted to use this to estimate the number of final submissions and hence the number of reviewers required. It helped to some extent, but many abstract submissions didn't turn into full paper submissions.

## Selection of Tracks

The submission form allowed authors to select the track in which their papers will be reviewed. Not all authors understood how to select the right track. We suggested that the keywords associated with each track be listed in the Call for Papers (similar to what was done in ACL 2018) and that the authors should refer to them when deciding which track their paper(s) should be submitted to. Given the very large number of submissions, we found it difficult to go through every submission to make sure that each one was submitted to the right track. In the end, we only managed to go through the submissions in the overloaded areas when attempting to load-balance the submissions across areas.

## Submission Day

The START system became heavily loaded just before the paper submission deadline, and we set the cutoff time in START later than the official deadline by **three hours**. The extra several hours was necessary to accommodate people who tried to submit on time but couldn't. Be prepared to receive many emails around the official submission deadline from authors with their PDF papers. Our suggestion is to ask them to try the system again rather than promising them that the PCs would upload the papers for them. There were just too many cases to handle manually.

## Desk Rejects

There was a lot of work to do after the paper submission deadline. PCs went through all papers, looking for problems that should lead to desk rejections, and identifying COIs. SACs checked for COIs, wrong areas, and obvious policy violations. We needed to move some submissions to suitable tracks and modify the submission type (long/short) for many submissions.

The typical problems leading to rejection without review included anonymity issues, format problems, length problems, and dual submissions.

## Review

### Review Form

We pretty much adopted the same review form as last year's EMNLP.

## Decision and Notification

We asked the SACs to make initial recommendations to us and then we made the final decisions for all the submissions. For SAC COI papers, we made the decisions directly. We identified the outlier papers which had high average scores but were recommended to reject, or had low average scores but were recommended to accept, and checked them carefully. For borderline papers, we considered both the rank of SACs and our own judgment to make final decisions.

After paper decision notification, the authors of a few (around 20) rejected submissions emailed us to argue with us about the decisions. We drafted a generic email to reply to them, explaining the details of the reviewing process.

Many of the emails involved complaining about the meta review not taking into account the author response. There were cases where discussion was needed but the AC didn't initiate it, and there were also numerous cases where the reviewers simply didn't look at the author response. In retrospect, more instructions should be given to ACs on when to conduct discussion among reviewers and how to write meta reviews (e.g., reviewer discussion should be summarized). Also, authors typically expect that the meta review would explain the decision process, but unfortunately in many cases, the final decision was made at a higher level. For instance, a paper that was recommended for acceptance by an AC was rejected later on by the



SAC and/or the program chairs when it was ranked against other papers in the same area. In other words, while authors expected to see the decision process in the meta review, it was not possible to provide all the information in the meta review. So the decision of whether to show the meta review to authors should also been taken seriously.