The Chair's Report on SIGMOD'10 Demonstration Program

Yufei Tao
Chinese University of Hong Kong
taoyf@cse.cuhk.edu.hk

1. STATISTICS AND FACTS

Program committee. The PC team[1] consisted of 4 and 8 reviewers from research labs and universities, respectively. Geographically, 7 reviewers were from North America, 3 from Europe and 2 from Asia.

Submission and review process. We had a total of 95 submissions, up from 86 last year. Each submission had no more than 4 pages, one more than last year. Unlike research-paper submissions (which were subject to double-blind reviewing), each demo paper had the full author list exposed. This tradition, which was followed in all the preceding SIGMOD conferences, is mainly because, in general, it is difficult, sometime even awkward, to achieve anonymity of systems, many of which are long-term projects with a homepage, and thus, is easily identifiable with Google.

The review process was quite standard. Each submission was refereed by three reviewers. Papers with unanimous positive (negative) ratings from all reviewers were immediately accepted (rejected). The rest (i.e., those with conflicts) then went through a discussion phase. The competition turned out to be rather heated. The high quality of the submissions rendered rejection a painful decision to make. We therefore decided not to limit the length of the acceptance list, but accept a paper as long as (i) its average rating was at least "weak accept", and (ii) it was championed by at least one reviewer. Eventually, 35 demos were accepted, giving an acceptance rate of 36.8%.

Demo program. The 35 accepted papers were grouped into 4 groups, each having 8 or 9 demos. At the conference, every group was allowed two sessions of 90 minutes (i.e., 3 hours in total) for demonstration to the conference attendees. Each group was provided with a table for equipment setup and an easel to support a poster. All sessions proceeded smoothly with a decent amount of visitor traffic.

[1]The full name list can be found at *http://www.sigmod2010.org/org_sigmod_demo_pc.shtml*.

2. BEST DEMO AWARD

The best-demo award has not been a tradition of SIGMOD. It was introduced for the first time in SIGMOD'05, but was suspended during 2006-2008, and only resumed last year. This year the award was given for the 3rd time. There seems to have been consensus that it is going to become a regular award.

Differences from the best-paper award. Deciding the winner of the best-demo award can be even more difficult than the best-paper counterpart. The main difficulty comes from the fact that the decision cannot be made by reading only the demo papers. Remember that a demo paper is short (4 pages this year). Furthermore, it does not help much even if they are allowed to be much longer, because the referees must actually see the system to get an accurate feeling about its power, user-friendliness, scalability, etc. This, in turn, immediately rules out the possibility of concluding the winner before the conference (as is possible, and usually practised, for best-paper awards).

The necessity of on-site evaluation creates more issues. The most serious of all is to find enough referees. Keep in mind that it is rather time consuming to assess a demo, because enough time must be allocated to enable a referee to weight the strengths and weaknesses of a system accurately. A rule of thumb is that the time of evaluating a demo should be comparable to the duration of a talk in a research session. Another fact to be taken into account is that a SIGMOD attendee usually considers it more important to attend research talks than seeing a demo. In other words, it is unrealistic to expect that a referee would agree to evaluate too many demos. On the other hand, to allow better global consistency in the evaluation, it is a good idea to have at least some dedicated referees that will inspect a large number

of the demos. These referees deserve respects, because they most likely are skipping many research talks that they would like to go to originally.

The second main issue is time. Besides the considerable amount of time needed for demo assessment, sufficient time must also be allocated for post-evaluation discussion, especially when several demos have close scores without an obvious tie-breaking solution. The time problem is particularly acute, if the demo chair hopes to announce the award at the award ceremony, which often takes place at slightly beyond the halfway of the conference. In that case, each demo may have only a single chance to be evaluated, because its second appearance would very probably be scheduled in the second half of the conference.

Referees. This year, I recruited two types of referees. A *global referee* is required to evaluate (at least) half of all the demos, whereas a *local referee* needs to assess 4-5 demos of one group (recall that each group has 8-9 demos, so a local referee effectively sees half of a group). A local referee needs to spend 90 minutes (i.e., one session), which appears to be a reasonable commitment. A global referee, however, is required to stay for 4 sessions, or totally, 6 hours.

There were 3 global referees: Chris Mayfield (Purdue Univ. USA), Andy Pavlo (Brown Univ. USA), and myself. The total number of local reviewers[2] was 17.

Evaluation process. All the demos were considered for the award according to the following 3 criteria:

- *System completeness.* We preferred full-fledged systems to half-baked ones. After all, the best demo paper should not be merely a "long version" of the experimental section of a research paper.
- *Problem importance.* Obviously, systems dealing with novel and urgent problems deserve additional credits.
- *Technical depth.* The award-winning system should have overcome certain major technical obstacles. This criterion rules out the "labor-intensive" systems built on either existing or straightforward techniques.

Each demo was assigned at least 1 global referee and 2 local referees. The assessment involved a *team phase* and, if time permitted, an *individual phase*. In the team phase, all referees assigned to a demo formed a team. The authors were given 18 minutes to impress the team, including up to 12 minutes of introduction to their system, followed by (at least) 6 minutes of questioning-asking by the referees. The value 18 was determined so that a referee who needed to evaluate 5 demos was able to do so within the 90-minute timeframe of a session. In the voluntary phase, the referees acted individually, and had the freedom of learning more about any demos of her/his choices. At the end of evaluation, the referees convened for a short discussion, and then scored the demo separately based on the 3 criteria mentioned earlier. The demo's final score is the average of the scores by all the referees. Finally, the global referees in their last (long) meeting ranked all the demos, and decided the winners.

All of the above information, including the names of the referees assigned to each demo, the assessment criteria, and the time allocation, etc., was announced to the demo authors well before the conference, and publicized online again well in advance.

Winners. The best-demo award eventually went to

A tool for configuring and visualizing database parameters,
by Vamsidhar Thummala, and Shivnath Babu,
from Duke University, USA.

We also decided to give honorable mentions to:

1. *DCUBE: Discrimination discovery in databases*, by Salvatore Ruggieri, Dino Pedreschi, and Franco Turini, from Universita di Pisa, Italy.
2. *K*SQL: A unifying engine for sequence patterns and XML*, by Barzan Mozafari, Kai Zeng, and Carlo Zaniolo, from UCLA, USA.

3. FOR THE FUTURE

Below are two issues that my experience in the past year has left with me. Unfortunately, they still confound me at this time. I hope the future demo chairs will be able to find clear answers, and by doing so, enhance the quality of their demo programs.

- *Systems before submission?* It came to my attention that, for some accepted demos, implementation was done after the papers were submitted. In other words, what was written in those papers was essentially in the authors' vision, as opposed to the real stuff. Should this be allowed?

[2]Name list on *http://www.cse.cuhk.edu.hk/∼taoyf/sigmod10bd.html*.

I exchanged some emails with Divy Agrawal (SIGMOD'10 PC chair of the research program) about this. Eventually, we decided to follow the policy of the past SIGMODs, which is not to discriminate against such submissions. After all, it is possible that, at the time of writing the paper, the authors came up with some nice ideas that they did not have time to implement yet, but they knew they would be able to accomplish. Not much harm could be done if they decided to write about these ideas in the paper first, just to meet the paper deadline. In any case, it appears to be rather difficult to clearly define how much implementation is appropriate at the time of submission, not to mention that it will be technically challenging for the committee to verify the degree of implementation during the review process.

On the negative side, one can argue that the absence of any implementation requirement makes it possible for some authors to say whatever they want at submission time to stand an unfairly high acceptance chance, and then implement only part of what they claim after the paper has been accepted. It is unlikely that anyone can find out at the conference anyway. Furthermore, even disregarding such a conspiracy theory, we should probably be reminded that a demo paper, by definition, ought to describe a demo — but how do you know what to demonstrate without even the system?

- *Video submission?* In SIGMOD'09, each demo submission could optionally be accompanied by a video submission. The idea also generated some interesting debates. One, in particular, revolves around the fact that the extra efforts of preparing a video actually do not necessarily increase the acceptance chance. This is in fact not too surprising. Without any video, a reviewer can only imagine the system. As dreams are often pretty, the reviewer may be picturing something fancier than the actual system, and thus, generously give a high score. A video therefore has the disadvantage of potentially smashing the beautiful imagination, and hence, incurring a lower score.

  As a side note, video submissions, when enforced in a compulsory manner, appear to be an effective weapon to discourage implementation after submissions.