SIGMOD 2016

**ACM SIGMOD Conference 2016**

June 26 - July 1, San Francisco, CA, USA

**Reviews For Paper**

| | |
|---|---|
| **Track** | Research, November 2015 Revision |
| **Paper ID** | 946 |
| **Title** | Generating Preview Tables for Entity Graphs |

**Masked Reviewer ID:** Assigned_Reviewer_1

**Review:**

| Question | |
|---|---|
| Overall Rating | Accept; an excellent paper you will fight to see accepted in SIGMOD |
| Comment on the ways in which the revision does and does not adequately address the issues and comments raised in the first round of reviews? (Feel free to comment both on the aspects of the revision relevant to your review and the other reviews.) | The authors seem to have addressed most of the concerns I had raised; supporting acceptance of this paper. |
| Does the revised manuscript address the comments and concerns you and the other reviewers raised during the first round of reviews? | Yes; the revision completely address all of the issues and comments raised by the reviews. |
| Detailed comments. Provide feedback about aspects of the paper that still need work or that should be revised (if the paper is accepted), or provide detailed feedback about why you feel the paper should not be accepted. | N/A |

**Masked Reviewer ID:** Assigned_Reviewer_2

**Review:**

| Question | |
|---|---|
| Overall Rating | Reject; despite its merits, the paper lacks novelty or technical depth, or still needs substantial revision before it is ready for publication |
| Comment on the ways in which the revision does and does not adequately address the issues and comments raised in the first round of reviews? (Feel free to comment both on the aspects of the revision relevant to your review and the other reviews.) | The revision presents a user study (Section 6.3) that is more in line with what was the requested in the reviews. A panel of "experts" was used to generate preview schemas for multiple domains. A user-study was undertaken that evaluates the quality of the previews generated by the algorithms in the paper against the expert-developed, freebase baseline, and closest-competitive alternate approach. The expert-generated schema seems to have been created in a reasonable process -- independent, deliberative process with the schemas being the same size as the target being compared against. The user study also seems to involve reasonable tasks (existential questions in the paper) and asks reasonable questions (Table 6). The analysis compares the time and accuracy in performing the tasks and also looks at overall satisfaction.<br><br>However, the analysis of the user study is a bit unsatisfactory. Specifically, first, it does not lead to focussed conclusions -- it establishes that preview generation is not simple (since expert and freebase schemas are not perfect). There is insufficient discussion as to how the auto-generated preview tables stack up the panel generated ones? Yes, there is a statement comparing the scores, but no analysis on the reasons. The fact that algorithmic preview tables score worse is acceptable (panel has more domain knowledge), but where is the justification that their score in the user study is a good/reasonable score?<br><br>Second, I believe the expert generated schemas should be used as the gold standard in 6.1.2 -- why bother comparing with Freebase, when you have a panel that generated preview tables? The study in 6.3 is there as a backstop to say that expert schemas are better than Freebase ones.<br><br>Third, there is still no analysis to separate out the need for the three variants of the algorithms? There has to be at least some discussion that says why one might prefer one over the other (possibly for different target applications) |
| Does the revised manuscript address the comments and concerns you and the other reviewers raised during the first round of reviews? | Mostly; the revision addresses most but not all of the issues and comments raised by the reviews. |
| Detailed comments. Provide feedback about aspects of the paper that still need work or that should be revised (if the paper is accepted), or | In view of the comments above, I believe there needs to be more analysis of the user study and a pointed discussion section that can lead readers to the conclusion that (1) preview tables are hard (I think this is shown), (2) preview tables are useful (not sure you show this, but it might be inferred from the results), (3) preview tables generated by the algorithms are reasonable (are you able to conclude this?), and (4) it is worth distinguishing between the variants presented (not shown).<br><br>You do not need to use all variants in the user study. The need for the variants is |

| | |
|---|---|
| provide detailed feedback about why you feel the paper should not be accepted. | orthogonal to the ability of the algorithms in general to come up with reasonable preview tables. So, the user study, for example, could just include the "tight" variable. There can be a separate analysis (or user-study) that only compares the three variants on specific task and then shows that they lead to qualitatively different results that are better suited for different tasks. |

**Masked Reviewer ID:** Assigned_Reviewer_3

**Review:**

| Question | |
|---|---|
| Overall Rating | Accept; an excellent paper you will fight to see accepted in SIGMOD |
| Comment on the ways in which the revision does and does not adequately address the issues and comments raised in the first round of reviews? (Feel free to comment both on the aspects of the revision relevant to your review and the other reviews.) | The reviewers addressed my main concerns in the revised version. However, I would like to see a discussion about the size of the schema used in the experiments as I mentioned in my first round of review (Weak Point 2). |
| Does the revised manuscript address the comments and concerns you and the other reviewers raised during the first round of reviews? | Mostly; the revision addresses most but not all of the issues and comments raised by the reviews. |
| Detailed comments. Provide feedback about aspects of the paper that still need work or that should be revised (if the paper is accepted), or provide detailed feedback about why you feel the paper should not be accepted. | As I mentioned before, the reviewers addressed my main concerns. If the paper is accepted, I would like to see a short discussion regarding the size of the schema used in the experiments and whether it is sufficiently big enough. |