First reviewer's review:

Evaluation <<<

- There weren't enough examples in the paper to give the reader a rich

understanding of the kinds of constraints that exist in practice. Table 1

provides the four kinds of similarity computations that were used in the paper,

but it would have been helpful to see examples supporting the rationale of each

of these. In general I would have liked to have seen several examples of actual

constraints upfront in the paper. The reason this is important for this

particular paper, is that the authors are basically forming (informal)

hypotheses that you can identify constraints based on similarities between

feature descriptions - and furthermore, there are different ways of computing

these similarities (i.e. full descriptions, objects, adj. modifiers etc).

Later in the paper the authors provide an example of a constraint that text

message notifications excludes the use of XML. As a reader, I would really have

liked to see rich examples of such constraints - so that I could judge whether

the ‘hypotheses' seem reasonable.

- It is unclear to me in the graph, how much feedback was included for the LU

values. Your experiment shows LU results, but doesn't explain the amount of

feedback used to achieve these results.

- Need to better explain phenomena such as why inner domain precision is worse

than cross-domain for GPL, and better for Weather Station. Given these unusual

results, I would have expected the authors to include more than two systems in

a cross-domain experiment. This is really not enough to draw any useful

conclusions in a general sense.

\*=--=\*=--=\*=--=\*=--=\*=--=\*=--=\*=--=\*=--=\*=--=\*=--=\*=--=\*=--=\*=--=\*=--=\*=--=\*

Second reviewer's review:

Weaknesses:

I enjoyed reading the paper and I was unable to identify major weaknesses.

However, some ideas to improve the paper are as follows:

- Section 2 is too short and could be merged into the introduction, which by

itself introduces part of the terminology. Alternatively, you could introduce

new terms as introduction of Section 3.

- The paper would profit from a discussion section, where you summarize the

main findings, and discuss the approach and its usefulness in practice, e.g.:

-- How to integrate the approach into RE tools and processes?

-- What are other influencing factors of the precision?

-- Does the interpretation/semantics of the constraints (in particular

excludes) only depend on the domain? I think it is rather dependent of the

organization (shared understanding) or the single people creating these

documents?

-- The treats section can be a subsection of the discussion.

- Table 1 is not easy to understand at the first glance. I would introduce

Figure 3 before or add a column "example".