Softw Syst Model (2008) 7:253–254
DOI 10.1007/s10270-008-0096-x

EDITORIAL

Fair treatment of evaluations in reviews

Robert France

Published online: 21 June 2008

As scientists, we understand and appreciate the value of evaluating the results of our research effort. As software engineers, we are painfully aware of the difficult challenges we must address when attempting to rigorously evaluate the methods, techniques, tools, languages, and other artifacts that we produce. The pressing problems that we tackle in the software and system modeling research domain can be classified as "wicked problems": we learn more about the nature of the problems we tackle through experimentation with proposed solutions. Rigorous evaluation of these solutions invariably entails costly and lengthy experimentation in industrial contexts. Experiments that seek to evaluate solutions based on novel or radically different ideas are particularly difficult to sell to potential industrial partners because the risks are not well-understood by all involved. Even with committed industrial partners, the wide variations in industrial development environments makes it difficult (if not foolhardy) to extrapolate the results beyond the specific industries. Despite the difficulties, there is no getting away from the reality that evaluation is key to developing progressively better solutions to wicked problems. As researchers, we must evaluate the products of our research. The responsibilities of manuscript authors with respect to the evaluation content are not the focus of this editorial; there are many published high quality articles on this topic. Rather, this editorial focuses on the responsibilities of reviewers when it comes to commenting on the evaluation content of submitted journal papers.

Many of us have seen or written manuscript reviews in which the lack of rigorous evaluation is used as a basis for rejecting a paper. In some cases this may be justifiable, but there are cases where the judgment is too harsh and could help stifle a promising idea in its early stages. As reviewers we should all be aware of the difficulty of doing rigorous evaluation in our domain, and, thus it is not enough to simply state in a review that "more evaluation is needed" or that "the evaluation is lacking". If a reviewer finds inadequacies in the evaluation content of a manuscript, he should be willing to state his expectations for a good evaluation of the work. A truly helpful review will point out questions that need to be answered by experiments and provide pointers to suitable evaluation tools. A good review should state the expectations of the reviewer with respect to the type and nature of the evaluation that is desired, and state clearly why the paper under review missed the mark. Reviewers should also appreciate that in some cases qualitative, well-reasoned evaluations is the best that can be reasonably done.

When stating expectations for evaluations, reviewers should ensure that they are reasonable with respect to the manuscript's content. For example, given a paper that proposes a new method for establishing and maintaining a variety of traceability relationships across models, it may be unreasonable to state the following in a review: "The authors need to perform an evaluation in industry to assess scalability before this work can be published". The problem with this expectation is that there are very few industries out there with a sizable repository of models to make such studies possible. So how do we as reviewers determine what is or what is not a reasonable expectation for an evaluation? It is not always easy to determine this, but one way is to ask ourselves what we would do given resources that are readily available to the community. The point about widely available resources is important, for example, if we have industrial partners that are willing to work with us, we should not assume that the authors are as fortunate.

Determining the adequacy of the evaluation content of a manuscript can be highly subjective in some cases. Authors

R. France (✉)
Colorado State University, Fort Collins, CO, USA
e-mail: france@cs.colostate.edu

present evaluations of their research results primarily to convince readers that the results are worthy. What convinces one reviewer may not convince another. It is thus important that a reviewer that is not convinced state clearly the questions they have that are not adequately answered in the evaluation.

While we focus on reviewers in this editorial, here are some words of advice for future SoSyM submitters. Your chances of getting a paper accepted are enhanced if you include at least a qualitative evaluation of the reported results.

In this issue

In this issue we are pleased to present a special section consisting of extended versions of the best papers accepted at SEFM 2005, and four papers in the regular section. The editorial for the special section presents an overview of the papers in this section.

In the regular paper "A Methodology for the Selection of Requirements Engineering Techniques" the authors Li Jiang, Armin Eberlein, Behrouz H. Far, and Majid Mousavi propose a method called Methodology for Requirements Engineering Technique Selection (MRETS), and describe how it can be used to select techniques that yield the highest quality requirements at a low cost. MRETS matches properties of techniques to properties of projects. It has been validated in an industrial project.

The regular paper titled "MDA Tool Components: A Proposal for Packaging Know-how in Model Driven Development" by the authors Reda Bendraou, Philippe Desfray, Marie-Pierre Gervais, and Alexis Muller proposes an approach that aims to improve the reusability of MDA artifacts (e.g., metamodels, UML profiles, model transformations) in different MDA tools. The paper introduces the concept of a MDA Tool Component that homogenizes the processes of packing, delivering, and executing MDA elements in MDA Frameworks. The authors define the types of artifacts to be included in a MDA Tool Component based on a set of metamodels.

The third regular paper is titled "Reducing Accidental Complexity in Domain Models" and is authored by Thomas Kühne and Colin Atkinson. In this paper the authors discuss the importance of reducing accidental complexity in modeling techniques, that is, complexity arising from a mismatch between the target problems and the techniques used to tackle the problems. They then discuss how accidental complexities arise in domain models and discuss ways of avoiding these complexities.

In the fourth regular paper "Improving the Accuracy of UML Metamodel Extensions by Introducing Induced Associations" by Xavier Burgués, Xavier Franch, and Maria Josep Ribó, the notion of induced associations is discussed. The paper discusses the need to distinguish M2 meta-associations that induce associations at the M1 level from those that do not. The paper also presents an approach for enforcing associations induced by meta-associations.