BP-Diff: A Tool for Behavioral Comparison of Business Process Models

Abel Armas-Cervantes[1], Paolo Baldan[2], Marlon Dumas[1], and Luciano García-Bañuelos[1]

[1] Institute of Computer Science, University of Tartu, Estonia
{abel.armas,marlon.dumas,luciano.garcia}@ut.ee
[2] Department of Mathematics, University of Padova, Italy.
baldan@math.unipd.it

Abstract. BP-Diff is a tool for identifying and diagnosing behavioral differences between pairs of business process models. BP-Diff identifies behavioral discrepancies involving pairs of tasks and provides both verbal and visual feedback to help users understand each discrepancy. The verbal feedback explains how a given pair of tasks is related in one model in contrast to the other model. Meanwhile, the visual feedback allows users to pinpoint the exact state where the discrepancy occurs. Unlike existing techniques, BP-Diff abstracts away from syntactical differences, focusing instead on behavior.

1 Introduction & background

The consolidation of business process models is, among others, a scenario where analysts need to have a clear understanding of the differences between business processes. Although, it is well-known that a pair of structurally different process models can represent the same behavior. In this regard, we address the problem of comparing process models based on their behavior. In short, we seek to determine if a pair of process models are behaviorally equivalent and if not, then we seek to produce feedback, as intuitive and as easy to grasp as possible, expressing the existing differences.

The foundations of the tool presented in this paper is the process model differencing technique introduced in [1]. Roughly speaking, given a process model, an abstract representation of the underlying behavior is computed. Such representation is an Asymmetric Event Structure (AES) [2], which is a model of concurrency that depicts the behavior of a process with two elements: events (instances of activities) and two binary behavioral relations. AES is a very expressive formalism and the same behavior can be represented with multiple non-isomorphic AESs. We have developed a method, described in a series of publications [1, 3, 4], to derive canonical AESs for process models for the purpose of comparison.

An AES can be seen as a full graph where the events are nodes and the relation between any pair of events is a labelled edge, i.e., the type of the relation is the label of the edge. Then, using this representation to compare a pair of business process models, one can compute canonical AES representing the behavior of processes and check graph isomorphism. In this regard, when a pair of AES are isomorphic then it implies that they represent equivalent behavior (in [1], we adopt an equivalence in the true concurrency spectrum). Conversely, if the representations are not isomorphic, then the AES allows to express the discrepancies between a pair of processes, possibly in its most basic form, as mismatching binary behavioral relations between tasks.

Oftentimes it is not the possible to establish a one-to-one correspondence between the events in a pair of AES, since an AES can contain more than a single event with the same label, or because one of the AES represents more behavior than the other. Thus, more sophisticated techniques for the comparison of a pair of AES are required in order to deal with general case. More specifically, given a pair of AES to compare, the technique in [1] proposes the use of error-tolerant graph matching techniques for finding a possible optimal mapping between the events of the event structures. Then, once the mapping has been computed, the comparison follows as before, I.e., testing for the graph isomorphism between a pair of AES, and outputting the feedback when discrepancies (mismatching relations) are detected.

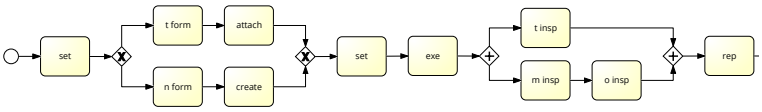

Fig. 1. Process model

Generally speaking, a single pair of activities can be in different behavioral relations depending on the run when they occur (what we call context). For example, in the process depicted in Figure 1 there is a run where the task *n form* precedes task *exe*; conversely, there is another run where *n form* does not occur together with *exe* (i.e., when *t form* occurs). In the technique presented in [1], we propose to give an approximate context together with the textual explanation of the difference. In this case, the approximate context consists of the last event(s) that need to occur before the discrepancy arises. Ideally one would require the complete run (the list of all tasks that occurred before the discrepancy), although it can hinder on the understandability of the explanation.

2 Overview of the tool

The presented tool is partly the implementation of the process model differencing technique introduced in [1]. Therefore, given a pair of process models, the

tool i) computes the canonically reduced AES of the behavior of each of the processes, ii) finds a possible matching between the events of both representations, iii) compares the AES, and iv) outputs the pairs of mismatching relations (as human-readable sentences with an approximate context).

As an extension of the as-is differencing technique, the tool offers a visual representation of a given discrepancy. In this case, using a graphical representation, it is possible to represent the whole context (run) for a given discrepancy. We consider that the visual representation of the discrepancy can complement the textual explanations by easing the understanding, and not overloading the user with text. E.g., consider the process model in Figure 2, in this case we are spotting a run that lead to the execution of the activities *m insp* and *o insp*. We highlight the tasks of interest in red, whereas the executed tasks that lead to the execution of *m insp* and *o insp* are highlighted in green. The numbers attached to each of the highlighted elements represent the times the task was executed (it is of special interest when there are cycles in the process). Thus, in the example of Figure 2, it is easy to see that the numbers in the gray box are attached to the tasks that lead to the execution of *o insp*; whereas, those in green are attached to the tasks that lead to the execution of *m insp*.

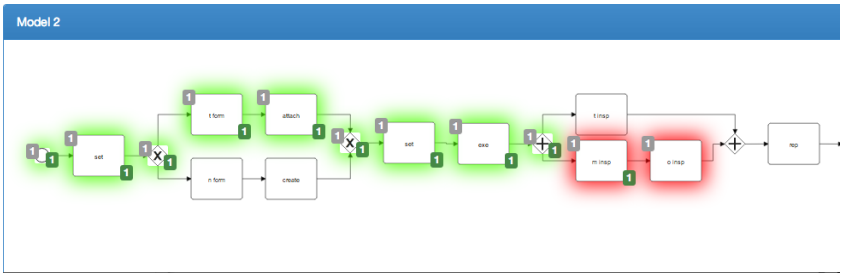

Fig. 2. Web interface

The tool provides a simple Web interface and it is depicted in Figure 3. On the top-left corner, the user can upload the pair of processes to be compared. Currently, the tool only support process models in BPMN modeling language. Once the process models have been submitted and the comparison has finished, then the textual descriptions of the encountered discrepancies are displayed on the left hand side of the screen. Finally, the models are rendered on the right-hand side of the window. Each of the discrepancies has a list of runs for both models, such that they show when the discrepancy occurs. Thus, the selection of any run associated to a discrepancy will produce a different coloring in the process. The tool relies on third-party libraries for the rendering of the models, i.e., Camunda BPMN JavaScript[3].

[3] https://github.com/camunda/camunda-bpmn.js

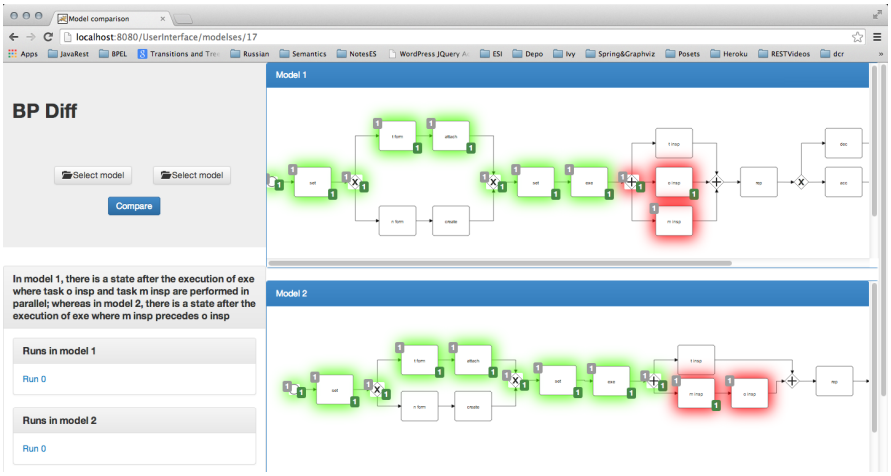

Fig. 3. Web interface

3 Maturity and significance

The tool has been tested in a relatively small set of business process models. The processes were modeled in Signavio [4]. As avenues for future research, we consider testing the tool with large process models to assess scalability. Secondly, we foresee an empirical usability evaluation of the diagnostics produced by our method with potential users.

The significance of this tool relies on the different applications where the diagnostics about the behavioral differences between pairs of business process models are required. For example:

1. Behavioral process model comparison: Determine if a pair of processes are equivalent and if not, provide a diagnostics of their differences. A variant of this problem is that of defining and calculating measures of behavioral similarity between pairs of processes.
2. Consolidation of multiple process variants into a single one: Provide accurate diagnosis about behavioral differences to guide analysis in the reconciliation of differences.
3. Compliance checking: Determine if a process is a behavioral refinement of the another.
4. Automated process discovery: Given a log, determine which binary relations exist in one graph.

To the best of our knowledge, BP-Diff is the first tool that provides both textual and graphical feedback about the behavioral differences found on pairs of business process models. The only tool we are aware of that implements behavior comparison is jBPT (available at `https://code.google.com/p/jbpt/`), which implements the theory of behavioral profiles [5]. As it is shown in [6], however, the diagnostics can be inaccurate. Moreover, the diagnostic generated with

[4] http://www.signavio.com/bpm-academic-initiative/

behavioral profiles cannot be directly translated into either textual or graphical feedback for human analysts.

The video showing a demo of the tool can be found in `http://math.ut.ee/~abela/bpdiffdemo/index.html`. Whereas, the tool is accessible in `http://diffbp-bpdiff.rhcloud.com/`.

References

1. Armas, A., Baldan, P., Dumas, M., García-Bañuelos, L.: Behavioral Comparison of Process Models Based on Canonically Reduced Event Structures. (to appear)
2. Baldan, P., Corradini, A., Montanari, U.: Contextual Petri Nets, Asymmetric Event Structures, and Processes. Information and Computation 2001 **171** 1–49
3. Armas, A., Baldan, P., García-Bañuelos, L.: Reduction of event structures under hp-bisimulation. Technical report `http://arxiv.org/abs/1403.7181`.
4. Armas-Cervantes, A., García-Bañuelos, L., Dumas, M.: Event structures as a foundation for process model differencing, part 1: Acyclic processes. In Beek, M., Lohmann, N., eds.: Web Services and Formal Methods. Volume 7843 of Lecture Notes in Computer Science. Springer Berlin Heidelberg (2013) 69–86
5. Weidlich, M., Mendling, J., Weske, M.: A foundational approach for managing process variability. In Mouratidis, H., Rolland, C., eds.: Advanced Information Systems Engineering. Volume 6741 of Lecture Notes in Computer Science. Springer Berlin Heidelberg (2011) 267–282
6. Armas-Cervantes, A., Dumas, M., García-Bañuelos, L., Polyvyanyy, A.: On the suitability of generalized behavioral profiles for process model comparison. In: WSFM. (2014) Accepted.