Information Systems 56 (2016) 304–325

Contents lists available at ScienceDirect

Information Systems

journal homepage: www.elsevier.com/locate/infosys

Diagnosing behavioral differences between business process models: An approach based on event structures

Abel Armas-Cervantes [*,a], Paolo Baldan [b], Marlon Dumas [a], Luciano Garcia-Bañuelos [a]

[a] Institute of Computer Science, University of Tartu, Estonia
[b] Department of Mathematics, University of Padova, Italy

ARTICLE INFO

Article history:
Received 18 December 2014
Received in revised form
16 September 2015
Accepted 21 September 2015
Available online 19 October 2015

Keywords:
Process model comparison
Asymmetric event structures

ABSTRACT

Companies operating in multiple markets or segments often need to manage multiple variants of the same business process. Such multiplicity may stem for example from distinct products, different types of customers or regulatory differences across countries in which the companies operate. During the management of these processes, analysts need to compare models of multiple process variants in order to identify opportunities for standardization or to understand performance differences across variants. To support this comparison, this paper proposes a technique for diagnosing behavioral differences between process models. Given two process models, it determines if they are behaviorally equivalent, and if not, it describes their differences in terms of behavioral relations – like causal dependencies or conflicts – that hold in one model but not in the other. The technique is based on a translation from process models to event structures, a formalism that describes the behavior as a collection of events (task instances) connected by binary behavioral relations. A naïve version of this translation suffers from two limitations. First, it produces redundant difference statements because an event structure describing a process may contain unnecessary event duplications. Second, this translation is not directly applicable to process models with cycles as the corresponding event structure is infinite. To tackle the first issue, the paper proposes a technique for reducing the number of events in an event structure while preserving the behavior. For the second issue, relying on the theory of complete unfolding prefixes, the paper shows how to construct a finite prefix of the unfolding of a possibly cyclic process model where all possible causes of every activity is represented. Additionally, activities that can occur multiple times in an execution of the process are distinguished from those that can occur at most once. The finite prefix thus enables the diagnosis of behavioral differences in terms of activity repetition and causal relations that hold in one model but not in the other. The method is implemented as a prototype that takes as input process models in the Business Process Model and Notation (BPMN) and produces difference statements in natural language. Differences can also be graphically overlaid on the process models.

1. Introduction

Large organizations often need to manage multiple variants of the same business process. For example, an order-

* Corresponding author.
E-mail addresses: abel.armas@ut.ee (A. Armas-Cervantes), baldan@math.unipd.it (P. Baldan), marlon.dumas@ut.ee (M. Dumas), luciano.garcia@ut.ee (L. Garcia-Bañuelos).

http://dx.doi.org/10.1016/j.is.2015.09.009
0306-4379/

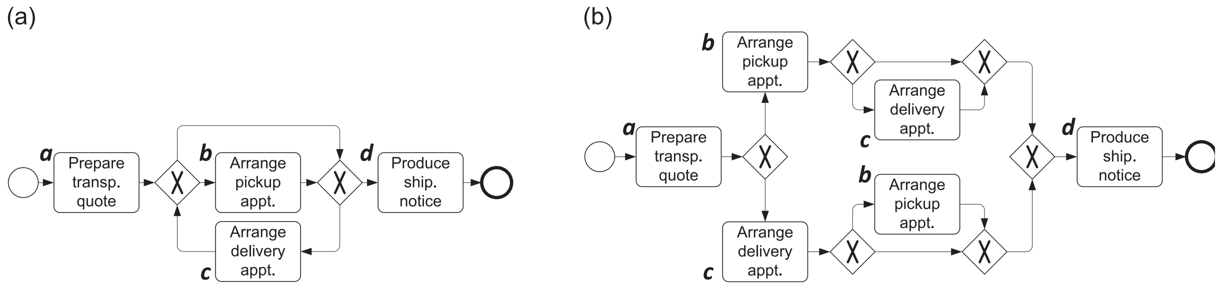

Fig. 1. Variants of business process models (a) M_1 and (b) M_2.

to-cash process may exist in multiple variants, corresponding to different products, different types of customers, different markets in which the company operates, or idiosyncratic choices made by multiple business units over time. During the ongoing management of these processes, analysts need to compare models of multiple process variants [1] in order to identify opportunities for standardization or to understand relative performance differences across variants.

Existing process model comparison methods can be classified into those based on the structure of the models and those based on their behavior. In some cases, a structural comparison – where nodes and/or edges are matched based on the topology of the model – is sufficient to understand the differences between two variants. However, two process models may be structurally different, yet behaviorally equivalent or they may be very similar structurally yet quite different behaviorally, as changes in a few gateways or edges might entail significant behavioral differences.

In this setting, this paper faces the problem of diagnosing behavioral differences between business process models. The paper presents a method to describe differences in terms of binary behavioral relations and activity repetition observed in one process but not in the other. We specifically deal with three elementary types of behavioral relations that, together with repetition, have been postulated as basic control-flow workflow patterns [2], namely causal precedence (corresponding to "sequence" in a process model), conflict (exclusive branches in a process model), and concurrency (parallel branches in a process model). For example, consider the models in BPMN notation in Fig. 1, describing an order fulfillment process, as presented in [3]. We aim at describing their differences via statements of the form: "*In M_1, there is a state after* Prepare transportation quote where Arrange delivery appointment can occur before Produce shipment notice *or* Arrange delivery appointment *can be skipped, whereas in the matching state in M_2,* Arrange delivery appointment *always occurs before* Produce shipment notice", and "*In M_1 activity* Arrange delivery appointment *occurs* 0,1 *or more times, whereas in M_2 it occurs at most once*".

Throughout the paper we assume that the input process models are given as Petri nets. This design choice enables the application of the presented comparison technique to any process modeling language with a mapping to this formalism. For example, a transformation of a large subset of BPMN to Petri nets can be found in [4].[1] In addition to providing a language-neutral representation, the use of Petri nets allows us to reuse a large body of existing theoretical results, for example the theory of unfolding [6,7].

Given that we focus on describing differences in terms of causality, conflict and concurrency relations, we adopt *event structures* [7] as an abstract representation of processes that explicitly recognizes these three types of relations. Event structures are a well-established model of concurrency where computations are represented as collections of events (activity occurrences) endowed with behavioral relations expressing dependencies between events. Various types of event structures have been introduced in the literature, comprising different binary behavioral relations, such as *prime event structures* [7] (PESs), where events are related by causal dependency and symmetric conflict, and *asymmetric event structures* [8] (AESs), where a form of asymmetric conflict between events is taken as primitive. A representation based on AESs can be more compact than one based on PESs. In fact, in the latter occurrences of the same activity in different contexts are necessarily represented as distinct events, leading to event duplication that is sometimes avoidable in AESs. For the purpose of comparison, more compact representations are desirable as they lead to more concise diagnoses of behavioral relations that exist in one process and not in the other. For this reason, the paper uses AESs as a basis for process model comparison.

In a prior work [9], we proposed a method for behavior-preserving reduction of AESs based on a *quotient operation*, which merges events corresponding to occurrences of the same activity in different contexts while preserving the overall behavior. However, the work in [9] shows that in some cases multiple non-isomorphic "minimal" AESs exist.

In this setting, the contribution of the paper is threefold: (i) we extend our previous work [9], by proposing a deterministic order on the quotient of an AES that leads to a uniquely determined minimal representation of a given acyclic process model; (ii) we propose a method for calculating an error-correcting (partial) synchronized product of two event structures from which differences can be diagnosed at the level of binary behavioral relations that hold in a state of a process model but not in the matching state of the other model; (iii) for cyclic process models, we rely on

[1] This transformation does not cover some BPMN constructs such as OR-joins, which cannot be straightforwardly translated into Petri nets [5].