A

Business Process Variability Modeling: A Survey

MARCELLO LA ROSA, Queensland University of Technology, Australia
WIL M.P. VAN DER AALST, Eindhoven University of Technology, The Netherlands
MARLON DUMAS, University of Tartu, Estonia and Queensland University of Technology, Australia
FREDRIK P. MILANI, University of Tartu, Estonia

It is common for organizations to maintain multiple variants of a given business process, such as multiple sales processes for different products or multiple bookkeeping processes for different countries. Conventional business process modeling languages do not explicitly support the representation of such families of process variants. This gap triggered significant research efforts over the past decade leading to an array of approaches to business process variability modeling. In general, each of these approaches extends a conventional process modeling language with constructs to capture customizable process models. A customizable process model represents a family of process variants in a way that a model of each variant can be derived by adding or deleting fragments according to customization options or according to a domain model. This survey draws up a systematic inventory of approaches to customizable process modeling and provides a comparative evaluation thereof with the aim of identifying common and differentiating modeling features, providing criteria for selecting among multiple approaches, and identifying gaps in the state of the art. The survey puts into evidence an abundance of customizable process modeling languages, which contrasts with a relative scarcity of available tool support and empirical comparative evaluations.

Categories and Subject Descriptors: H.4.1 [**Office Automation**]: Workflow management; A.1 [**Introductory and survey**]

General Terms: Design, Management, Standardization

Additional Key Words and Phrases: Variability modeling, process model, customizable process model

ACM Reference Format:
La Rosa, M., van der Aalst, W.M.P., Dumas, M. and Milani, F.P. 2015. Business Process Variability Modeling: A Survey. ACM Comput. Surv. V, N, Article A (January YYYY), 45 pages.
DOI = 10.1145/0000000.0000000 http://doi.acm.org/10.1145/0000000.0000000

1. INTRODUCTION

The co-existence of multiple variants of the same business process is a widespread phenomenon in contemporary organizations. As a concrete example, The Netherlands has around 430 municipalities, which in principle execute the same or a very similar set of processes. All municipalities have processes related to building permits, such as the process of handling applications for permits and the process for handling objections against such permits. Due to demographics and political choices though, each municipality executes its processes differently. Variations are justified by different priorities and customs, often referred to as the "Couleur Locale". At present, these differences have come to be accepted and there is no willingness to flatten them out. Still, capturing multiple municipality processes in a consolidated manner is necessary in order to develop information systems that can support multiple or all municipalities at once.

Authors' addresses: M. La Rosa, Queensland University of Technology, GPO Box 2434, Brisbane, QLD 4001, Australia; W.M.P. van der Aalst, Eindhoven University of Technology, PO Box 513 NL-5600 MB Eindhoven, The Netherlands; M. Dumas and F.P. Milani, University of Tartu, J. Liivi 2, Tartu 50409, Estonia.

DOI 10.1145/0000000.0000000 http://doi.acm.org/10.1145/0000000.0000000

Similarly, *Suncorp Group* – the largest insurance group in Australia – offers a range of insurance products, including home, motor, commercial and liability insurance. Each product exists for different brands of the group (e.g. Suncorp, AAMI, APIA, GIO and Vero). As a result, there are more than 30 variants of the process for handling an insurance claim at Suncorp Group. There is a case for modeling and maintaining these variants in a consolidated manner, not only to avoid redundancy, but also so that improvements and automation efforts made on one variant can benefit other variants.

The application of conventional business process modeling approaches [Mili et al. 2010] to families of process variants requires one of two paths to be chosen. Either each variant is modeled separately, resulting in duplication as the variants have much in common, or multiple variants are modeled together, leading to highly complex consolidated models, which hampers the analysis and maintenance of individual variants.

Motivated by this observation, a number of approaches to model families of business process variants have emerged. A common trait of these approaches is that they support the representation of a family of business process variants via a single model, from which each variant can be derived via certain model transformations. We use the term *customizable process model* to refer to such a consolidated model of process variants, and the term *variation point* to indicate an element of the customizable process model that can be customized via transformations.

A wide array of approaches to customizable process modeling have been proposed in recent years, without it being generally clear what tradeoffs they strike relative to each other and how should potential users select an approach for a given purpose. In this setting, this survey draws up a systematic inventory of approaches to customizable process modeling, identifies and classifies major approaches in the field and provides a comparative evaluation aimed at answering the following questions:

— **RQ1.** What are the commonalities and distinctive features of approaches to customizable process modeling?
— **RQ2.** What criteria can be used to select between different approaches?
— **RQ3.** What general limitations or research gaps exist in the literature on customizable process modeling that may require further work?

The rest of the article is organized as follows. Section 2 delimits the scope of the survey. Section 3 defines and justifies the criteria used to analyze approaches in the field. Section 4 presents a working example. Sections 5–9 illustrate and analyze major approaches in the field identified from a systematic literature review. Section 10 provides a synthesis of the commonalities and differences between the surveyed approaches and answers the research questions framed above. Section 11 positions this survey with respect to related work. Finally, Section 12 exposes common limitations, leading to an outline of future research directions.

Six appendixes complement the survey. Appendix A describes the literature search procedure and summarizes the results thereof. Appendix B surveys secondary approaches identified during the search process. Appendices C and D discuss techniques employed by the surveyed approaches to provide decision support during process customization and to ensure correctness of the customized process models. Finally, Appendix E provides a list of all relevant terms and their definitions used in this paper, while Appendix F shows a mapping between the different process modeling languages used to exemplify the approaches surveyed.

2. SCOPE

Customizable process models capture a family of process model variants in a way that the individual variants can be derived via transformations, e.g. adding or deleting fragments. Accordingly, a customizable process model encapsulates customization decisions between process variants that need to be made either at design-time or run-

time. *Design-time customization* decisions lead to a customized process model that is intended to be executed in a particular organizational setting. Hence, these decisions affect all instances of the customized process executed in this setting. The timeframe associated with these decisions may be long (e.g. months or years). In contrast, *run-time customization* decisions are punctual and affect only one or a few process instances. Such decisions may be visualized on top of a process model, but they are not intended to modify the executed process model itself, beyond its effects on the process instance(s) where the decision is applied.

Processes where customization decisions are made at run-time are called *flexible processes* [Reichert and Weber 2012]. The challenges associated with managing such processes have been widely studied in the literature [Rinderle et al. 2004; Weber et al. 2008]. The present survey focuses on *design-time process variability management* as opposed to run-time flexible process management. In other words, the focus is on capturing a family of processes via a single process model that is customized at design-time. Approaches to run-time flexible process management are generally not concerned with maintaining multiple process models that together form a family of processes. Instead, these approaches rely on a unitary process model. In some approaches, this unitary process model is seen as an indicative roadmap with respect to which individual process instances may deviate [Reichert and Dadam 1998], while in other approaches – e.g. Declare [Pesic et al. 2007] or Pockets of Flexibility [Sadiq et al. 2001; Sadiq et al. 2005]) – the process model is left underspecified and individual process instances refine this underspecified model rather than deviating from it. In both cases, there is still a single process model that serves as a reference during process execution.

Customization decisions may result in the removal or addition of behavior to a customizable process model. In this respect we distinguish two approaches to variability management: *by restriction* and *by extension*.

Variability by restriction starts with a customizable process model that contains all behavior of all process variants. Customization is achieved by restricting the behavior of the customizable process model. For example, activities may be skipped or blocked during customization. In this setting, one can think of the customizable process model as the union or Least Common Multiple (LCM) of all process variants. Customizable process models of this type are sometimes called *configurable process models*.

Variability by extension takes the opposite starting point. The customizable process model does not contain all possible behavior, instead it represents the most common behavior or the behavior that is shared by most process variants. At customization time, the model's behavior needs to be extended to serve a particular situation. For example, one may need to insert new activities in order to create a dedicated variant. In this setting, one can think of a customizable process model as the intersection or Greatest Common Denominator (GCD) of all process variants under consideration.

This survey covers *both* variability by restriction and by extension. In fact, as discussed later, it is possible for one same approach to combine both types of variability.

Customizable process models ought to be distinguished from so-called *reference process models* [Fettke and Loos 2003; Rosemann 2003]. Some vendors and consultancy firms provide reference process models that are intended to capture common knowledge or best practices in a given field (e.g. in supply chain management or IT service delivery). While a reference process model can be very useful in its own way, it should be understood that it is in essence a concrete process model intended to be used as an example. Reference process models do not support customization in a structured manner. In this survey, we focus on approaches that provide support for customization rather than serving only as reference.

Having discussed the scope of the survey, we next define as set of criteria to characterize the approaches in the field.

3. EVALUATION CRITERIA

To derive criteria for assessing approaches to business process variability modeling, we analyzed the solution space using the six "W questions" (Who, What, Where, When, Why and How). We determined that the "who" and "why" questions do not allow us to distinguish between approaches in the field, since all the approaches identified in the search (cf. Appendix A) have the same aim, i.e. to support process modelers ("who") in the definition of customizable process models and in the customization thereof ("why"). Similarly, the "where" question is not relevant as there is no spatial dimension that distinguishes approaches in the field. The "when" question ("when does customization occur?") has been discussed in the previous section (design-time vs. run-time) and a choice was made to focus on design-time, given that run-time customization has been studied as a separate topic in the literature (cf. process flexibility). This leaves us with the "what" and "how" questions, which we refine into: "What is captured in the customizable model?" and "How are customized models derived from customizable ones?".

To answer the "what" question, we reuse a classification of elements of a process model spelled out in previous surveys, whereby the elements of a process model are divided into those concerned with the *control-flow perspective* (the flow of control between activities), the *resource perspective* (organizational aspects), and the *object perspective* (physical and data objects manipulated in the process) [Georgakopoulos et al. 1995; Mili et al. 2010]. Accordingly, we characterize process variability modeling approaches depending on their support for each of these three perspectives.

Another classification of process models identified in previous work is based on their purpose [Georgakopoulos et al. 1995]. Along this direction, we distinguish between *conceptual* process models, which are intended for communication and analysis, and *executable* ones, which are intended for deployment in an execution engine.

Moving to the "how" question, we note that customized process models are derived from customizable models by applying transformations based on decisions made by a user. Thus, customization involves *decisions* and *transformations*.

The transformations applied during customization can be classified into those that restrict the process behavior captured by the customizable process model, e.g. by removing an element (customization by *restriction*), and those that extend the process behavior, e.g. by adding an element (customization by *extension*), as discussed in the previous section. Meanwhile, customization decisions can be expressed in terms of concepts that refer to the domain of discourse (abstract level), or concepts related to the process model itself (concrete level). Thus, approaches can be assessed depending on whether their customization decisions support *abstraction* to the domain level or not. Putting aside the concepts used to express decisions, some approaches guide the user step by step when making these decisions (i.e. by presenting the decisions in a certain order) and prevent inconsistent of irrelevant decisions to be made, while other approaches leave it up to the user to decide what decisions to perform and in what order. Accordingly, we can assess approaches depending on the *guidance* they provide during customization.

Transformations applied to derive a customized process model may in some cases lead to syntactically incorrect models, whether structurally or behaviorally incorrect. Some approaches guarantee that the customized process models are correct, but others do not. Accordingly, we can characterize an approach depending on whether or not it guarantees *structural* and/or *behavioral correctness* of the customized process model.

With respect to the "how" question, we considered alternative criteria. Since customizable process models generally extend a host modeling language, the latter could be used as a classification criterion. We did not retain this criterion because we observed that it is not a fundamental characteristic of an approach. An approach that has been designed for EPCs can be adapted to BPMN and vice-versa. We also consid-

ered the specific abstraction mechanism as a classification criterion. In this respect, approaches may differ in terms of the mechanism employed to link the elements of the process model to elements of the domain of discourse. Some approaches rely on simple "annotations" attached to model elements referring to implicitly defined elements of the domain of discourse, while others may opt for a more explicit linkage, where elements of the process model are linked to concepts in an explicit domain model (or vice-versa). The latter could be further subdivided depending on the approach employed to represent the domain model (e.g. feature model vs. questionnaire model). We opted however to simply classify approaches depending on whether they support abstraction or not, because we found that the choice of the domain modeling approach and the choice of the mechanism for linking the domain model to the process model are very approach-specific. In Section 5 we discuss these design choices for each approach separately.

Having identified assessment criteria based on "what" and "how" questions, we moved to the "meta" level, by considering the design of the approach itself. Research papers that propose customizable process modeling approaches rely, implicitly or explicitly, on a design science method [Hevner et al. 2004]. According to design science principles, the conceived artifacts should be specified, implemented where applicable, and validated to determine if they fulfill the intended requirements. Artifacts in the field under study can be specified informally or formally. They may or may not be implemented as a prototype. And they may or may not be validated in order to assess their applicability and qualities. Accordingly, we identify three extra-functional requirements: *formalization*, *implementation* and *validation*. The criteria resulting from the above analysis are explained below.

1 **Scope.** This category refers to the "what" question discussed above. It is decomposed into two sub-categories: *Process Perspective* and *Process Type*.
  1.1 **Process Perspective.** This category refers to the supported process modeling perspectives.
    1.1.1 **Control flow.** Ability of a customizable model to capture variability along the control-flow perspective, i.e. activities and routing elements such as gateways can become variation points (e.g. capturing that a credit history check is not required in some of the variants of a loan origination process). A language is considered to only partially fulfill this criterion if routing elements or activities are not customizable, or if such elements are customizable but the corresponding customization options are not graphically represented.
    1.1.2 **Resources.** Ability of a customizable model to capture variability in the involved human and non-human resources, i.e. resources can become variation points (e.g. capturing that a risk assessor is not involved in some of the variants of a claims handling process). A language partially fulfills this criterion if resources are customizable but the options are not graphically represented.
    1.1.3 **Objects.** Ability of a customizable model to capture variability in the physical and data objects produced and consumed by a process, i.e. objects can be become variation points (e.g. capturing that an invoice is not required in some of the variants of an order-to-cash process). A language partially fulfills this criterion if objects are customizable, but their customization options are not graphically represented.
  1.2 **Process Type.** This category refers to the purpose of the process models.
    1.2.1 **Conceptual.** An approach meets this criterion if it is designed to support conceptual process models only, i.e process models that are not

meant to be executed on top of a Business Process Management System (BPMS).

1.2.2 **Executable.** An approach is considered to fulfill this criterion if the customization prevents or resolves inconsistencies in the associations between activities and data objects, thus making the customized models executable on top of a concrete BPMS. If the customized models can be executed on a BPMS, but these inconsistencies are not addressed, the approach is considered to only partially fulfill the criterion. Similarly, the criterion is partially fulfilled if there is no BPMS that can support the execution of the customized models, even if inconsistencies are prevented or resolved by the approach.

2 **Customization Type.** Do the supported transformations restrict/extend the process behavior?

2.1 **Restriction.** An approach matches this criterion if a process model is customized by restricting its behavior.

2.2 **Extension.** An approach matches this criterion if a process model is customized by extending its behavior.

An approach could in principle support both criteria, i.e., there could be transformations to restrict some parts and extend others.

3 **Supporting Techniques.** This category refers to techniques to support the customization of process models. The two sub-categories are based on common functionality frequently reported in the literature: decision support for the selection of suitable customization options and ensuring the correctness of the customized model.

3.1 **Decision Support.** How are users supported in their customization decisions?

3.1.1 **Abstraction.** An approach supports process model abstraction if users can customize a model without directly referring to its model elements, but instead to properties of the application domain (e.g. customizing an order-to-cash process model based on the available sales channels, rather than based on the activities and gateways that are customizable in the model).

3.1.2 **Guidance.** This criterion is met if there is support to: (i) guide users when making customization decisions, e.g. in the form of recommendations for selecting one option or another; and (ii) prevent users from making inconsistent or irrelevant customization decisions from a domain viewpoint. Approaches that only provide support for one of these two aspects, partially fulfill this criterion.

3.2 **Correctness Support.** Is the syntactical correctness of the customized models guaranteed? Syntactical correctness is divided into correctness of the model structure, and correctness of the model behavior.

3.2.1 **Structural correctness.** Ability to guarantee the correct structure of the customized models, e.g., by avoiding disconnected nodes.

3.2.2 **Behavioral correctness.** Ability to guarantee the correct behavior of the customized models, e.g., by avoiding behavioral anomalies such as deadlocks and livelocks when the model is instantiated. In other words, the model must be *sound* [van der Aalst et al. 2011], i.e., it should always be possible to complete any process instance properly.

4 **Extra-Functional.** Criteria related to the design of the approach itself.

4.1 **Formalization.** Some approaches only present ideas and do not provide concrete algorithms or definitions. Therefore, we include a criterion indicating whether the approach has been described rigorously in terms of mathematical notations. In order to fulfill this criterion, the approach has to be formally

defined, including algorithms used during customization. If such algorithms are missing, the approach partially fulfills the criterion.

4.2 **Implementation.** Approaches may only exist on paper. However, the usability and maturity of an approach heavily depends on tool support to design and customize customizable process models. If the approach is fully implemented (including algorithms used during customization), then this criterion is fulfilled. Approaches with partial implementations, e.g. only offering design or customization support, partially fulfill this criterion.

4.3 **Validation.** The applicability of some approaches has been validated using real-life process variants and through discussions with domain experts, but this does not necessarily apply to all approaches. An approach fulfills this criterion if it has been tested on models not created by the authors, and the results verified by domain experts. If one of these two aspects is lacking (e.g. an approach that has been validated without the involvement of domain experts) then this criterion is only partially fulfilled.

The next section introduces an example of a family of process variants that is used later to illustrate the approaches retrieved by the search described in Appendix A.

4. ILLUSTRATIVE SCENARIO

The example process family described in this section is the result of a case study in picture post-production that we conducted with domain experts from the Australian Film, Television and Radio School (AFTRS) in Sydney.[1]

In the film industry, picture post-production (post-production for short), is the process that starts after the shooting has been completed, and deals with the creative editing of the motion picture. Figure 1 shows several variants of the picture post-production process. A process model is a directed graph consisting of nodes of type event, activity and gateway and arcs (called sequence flows) linking these elements. Events are triggers to and signal the results of activities, or of the entire process (e.g. a start event triggers the entire process, while an end event signals its completion). Activities capture work done in the process. Gateways are used to model alternative and parallel branching and merging and are divided into splits (with multiple outgoing flows and one incoming flow) and joins (with multiple incoming flows and one outgoing flow). Splits and joins have a logical type. They can be of type OR or XOR (for inclusive, resp., exclusive decision and merging) and AND (for parallelism and synchronization).

The example in Figure 1 is represented in the *Event-driven Process Chains (EPCs)* language [Davis and Brabander 2007]. There is a variety of languages, besides EPCs, to represent process models, e.g. BPMN, UML Activity Diagrams, YAWL, BPEL. While in this paper we will illustrate process model examples using different languages, depending on the approach being reviewed, for uniformity we will always use the terminology described above, which is borrowed from the BPMN standard, and abstract from language-specific terms. Appendix F provides a mapping between the languages used to exemplify the approaches surveyed.

As depicted in Figure 1, post-production starts with the receipt, from the shooting that needs to be prepared for editing. The footage can either be prepared on film (see e.g. variant a of Figure 1, where activity "Prepare film for editing" in performed), on tape (e.g. variant b, where activity "Prepare film for editing" is performed) or on both media (variant d) depending on whether the motion picture was shot on a film roll and/or on a tape. Next, the medium is edited offline to achieve the first rough cut (thus activity "Edit offline" exists in all variants). However, after this, an online editing is carried out if the footage was shot on tape (variants b and c), while a negmatching is

[1] See `www.aftrs.edu.au`.

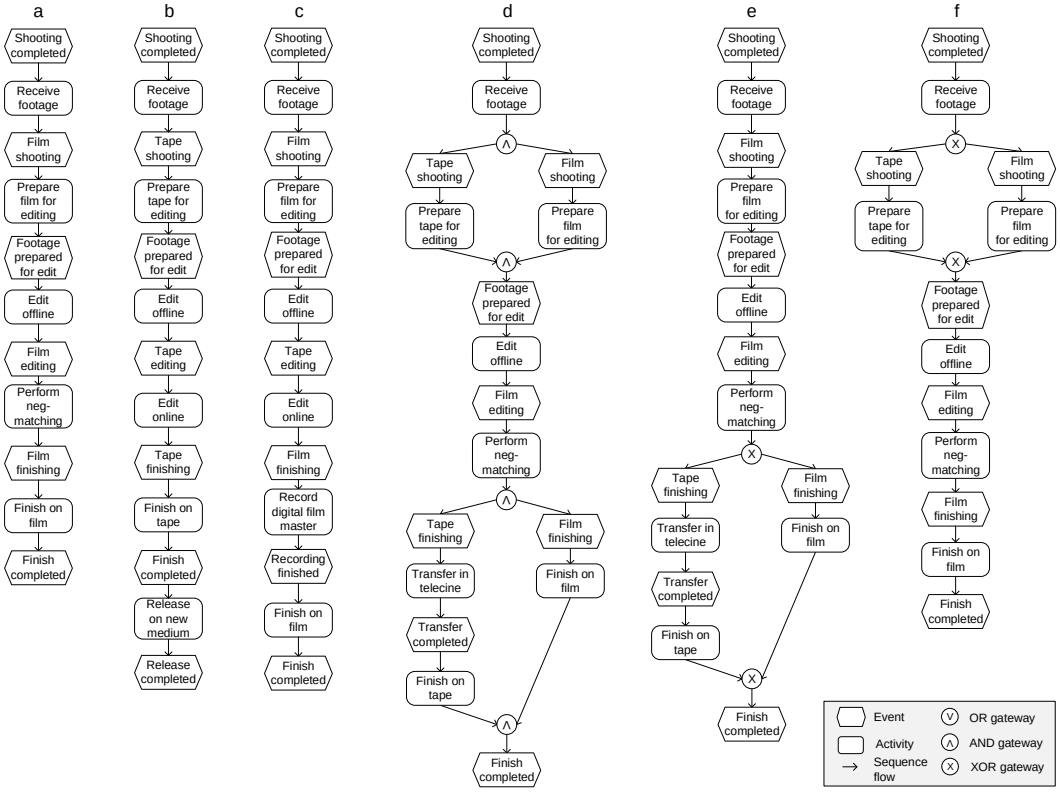

Fig. 1: Different variants of the picture post-production process in the EPC language.

performed if the footage was shot on film (e.g. variant a). Online editing is a cheap editing procedure suited for low-budget movies typically shot on tape. Negmatching offers better quality results but entails higher costs; thus it is more suitable for high-budget productions typically shot on film. The choice between online editing and negmatching is an important post-production decision: depending on drivers such as budget, creativity and type of project, one option, the other or both need to be taken. Thus, each variant in Figure 1 reflects a common practice in post-production. For example, variant a is a typical low-budget practice (shooting and releasing on tape), whereas variant d illustrates a more expensive procedure (shooting and releasing on both tape and film).

The final step of post-production is the finishing of the edited picture. This can be be done on film (see variant a), on tape (variant b) or on both media (variant d). The finishing may involve further activities based on the combination of editing type and final medium. For example, if the editing was done online and the final version is on film, a digital film master is to be recorded from the edited tape (see variant c). Alternatively, if a negmatching was performed and the final version is on tape, the edited film is to be transferred onto a tape via a telecine machine (variants d and e).

The process may conclude with an optional release on a new medium (e.g. DVD or digital stream), which follows the finishing on tape or film (for example, in variant b the release on new medium follows a tape finish).

5. OVERVIEW OF PROCESS MODEL CUSTOMIZATION APPROACHES

We conducted a literature search using the protocol described in Appendix A. This search resulted in 66 relevant publications. In many cases, multiple publications pertain to the same approach. Also, some approaches are subsumed by other approaches, i.e., the concepts in one approach are contained in another. By grouping the publications accordingly, we found that the 66 publications cover 23 approaches, out of which 11 main approaches subsume the other 12 approaches.

The 66 publications covered by this survey are listed in a supplemental spreadsheet available at `https://goo.gl/mmxZf3`. For each approach, the table identifies a primary (earliest) publication describing the approach and where available, additional publications describing further aspects of the same approach.

A histogram of papers per year of publication is shown in Figure 2. This histogram is based on the list of publications satisfying the inclusion and exclusion criteria defined in the search protocol (cf. Appendix A).[2] The histogram shows an increasing trend of publications on the topic starting in 2005 and reaching a peak in 2010.

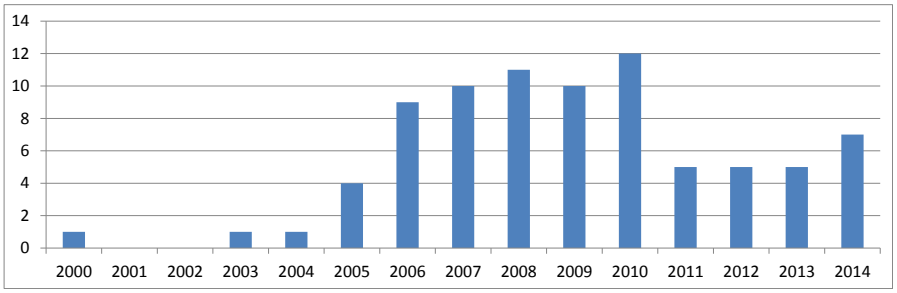

Fig. 2: Number of publications on process model variability management via customizable process models.

We classified the 23 identified approaches by asking the following question for each approach: "How does the approach capture the relation between an element or set of elements of a customizable process model, and a corresponding element or set of elements of each of the possible customized process models thereof?"

Answering the above question for each approach led us to observe that in some approaches a node of the customizable model can be retained, removed or its behavior can be restricted, by selecting one of multiple possible customization options. This class of approaches is hereby called *node configuration*. In other approaches, an element (i.e. a node or a sequence flow) of the customizable process model is linked to a predicate over a domain model via an *annotation*. Customization then takes place by evaluating these predicates with respect to an instantiation of the domain model. We call this class of approaches *element annotation*. In a third class of approaches, a given activity in the customizable process model can be replaced by one of multiple specialized versions thereof. These approaches only allow specialization of activities and their attributes, and not of other types of elements. Hence, we call this class *activity specialization*. Finally, in a fourth type of approaches, the relation between the customizable process model and its customized models is specified by means of *change operations* that can add, delete or modify fragments of the customized model. Since the latter approaches can manipulate entire fragments, we call this class *fragment customization*.

[2]The histogram includes papers with less than 10 citations, even if this was an exclusion criterion, since its intent is to show the volume of research publications in the field over time. Hence, the number of references covered by the histogram is larger than the 66 publications mentioned above.

The taxonomy induced by the above observations is given in Figure 3, where the main approaches are shown in bold, and the subsumed ones are listed under their respective main approach. In line with previous work on variability modeling [Svahnberg et al. 2005; Bachmann and Clements 2005; Becker et al. 2007a], we use the term *variability mechanism* to refer to a set of modeling constructs and their corresponding semantics, used to specify the relations between a customizable process model and its possible customized models (a list of all terms and their definitions is provided in Appendix E). The taxonomy presented in Figure 3 effectively classifies the variability mechanisms underpinning the identified approaches.

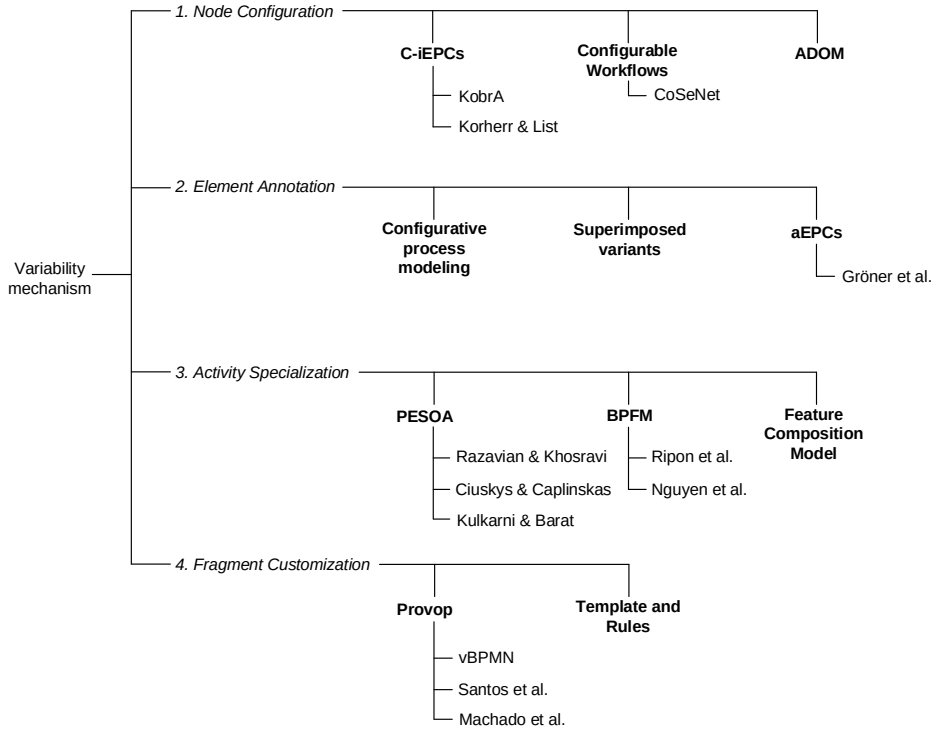

Fig. 3: Taxonomy of approaches for process model customization.

In the next four sections, we briefly introduce and evaluate the main approaches under each of the groups identified above, using the 14 criteria described in Section 3. We discuss the subsumed approaches in Appendix B. The results of the assessment are then summarized in Table XII (Section 10). The evaluation is complemented by an overview of techniques for customization decision support (Appendix C) and for correctness support (Appendix D).

The assessment of each approach was performed independently by two authors of this paper. The results were compared in order to resolve inconsistencies with the mediation of a third author. Finally, we sought confirmation of our assessment from the authors of each primary publication.

6. GROUP 1: NODE CONFIGURATION

In the approaches in this group, a variation point is a node (called *configurable node*) of the customizable process model which is assigned different customization options.

Activities, events, gateways, as well as resources and objects associated with activities, may be marked as configurable nodes. Customization is achieved by selecting one customization option per configurable node. Each configurable node has an option to keep the node as is in the customized model, and one or more options to restrict its behavior.

Configurable activities, events, resources and objects can be customized by being kept *on* (they remain in the customized model), or switched *off* (they do not appear in the customized model). The semantics of switching an activity or event off is approach-specific, i.e. the activity or event may be hidden without breaking the path to which it belonged, or be removed altogether. Configurable gateways can be customized to an equal or more restrictive gateway, in such a way that the customized process model produces the same or fewer execution traces than the customizable process model.

Three main approaches fall into this group: C-iEPCs, Configurable Workflows and ADOM. They support different subsets of the above configurable node types and customization options, e.g. C-iEPCs do not support configurable events. In addition to customization by restriction, ADOM offers a weak form of extension, in that extension points are not identified in the model.

6.1. Configurable integrated Event-driven Process Chains (C-iEPCs)

Configurable integrated EPCs (C-iEPCs) [Rosemann and van der Aalst 2003; Dreiling et al. 2005; Dreiling et al. 2006; La Rosa et al. 2011] are an extension of the EPC language. Essentially, an iEPC is an EPC with resources and objects assigned to activities. A C-iEPC model is intended to capture the least common multiple of a family of iEPC variants. Differences among the various process variants are indicated by configurable nodes. Each configurable node can be assigned a set of customization options, each referring to one or more process variant. Customization is achieved by restricting the behavior of the C-iEPC by assigning one customization option to each configurable node. Then the C-iEPC is transformed into an iEPC by removing all those options that are no longer relevant. By doing so, one can derive one of the original variants of the given process family.

Activities and gateways can be marked as configurable with a thicker border. Events cannot be customized. Figure 4 shows the C-iEPC model for the post-production example, which captures all variants of Figure 1.

Configurable gateways can be customized to an equal or more restrictive gateway. A configurable OR can be left as a regular OR (no restriction), or restricted to an XOR or to an AND gateway. Moreover, the number of its outgoing flows (if the gateway is a split), or the number of its incoming flows (if a join), can be restricted to any combination (e.g. two flows out of three), including being restricted to a single flow, in which case the gateway disappears.

For example, we can capture the choice of the shooting medium by customizing the first OR-split in Figure 4. We can restrict this gateway to the outgoing flow leading to event “Tape shooting” if the choice is tape. As a result the branch starting with event “Film shooting” is removed, and vice versa. Restricting the gateway to an AND-split ensures that both media are prepared for editing. In the three cases above, we anticipate the decision of the medium at configuration-time. Alternatively, by configuring this gateway to an (X)OR-split, we postpone the decision till run-time, when the post-production process is actually enacted (see e.g. variant f in Figure 1).

Configurable activities can be kept *on* or switched *off*. In the latter case, the activity is simply hidden in the customized model. In addition, they can be customized to *optional*. This allows the deferral of the choice of whether to keep the activity or not until run-time. For example, function “Release on new medium” is configurable in Figure 4, so we can switch it off for those post-production projects where this is not required.

Resources (called *roles* in C-iEPCs) and objects can also be made configurable. In the C-iEPC semantics, when an object is used as input to an activity, it can be marked

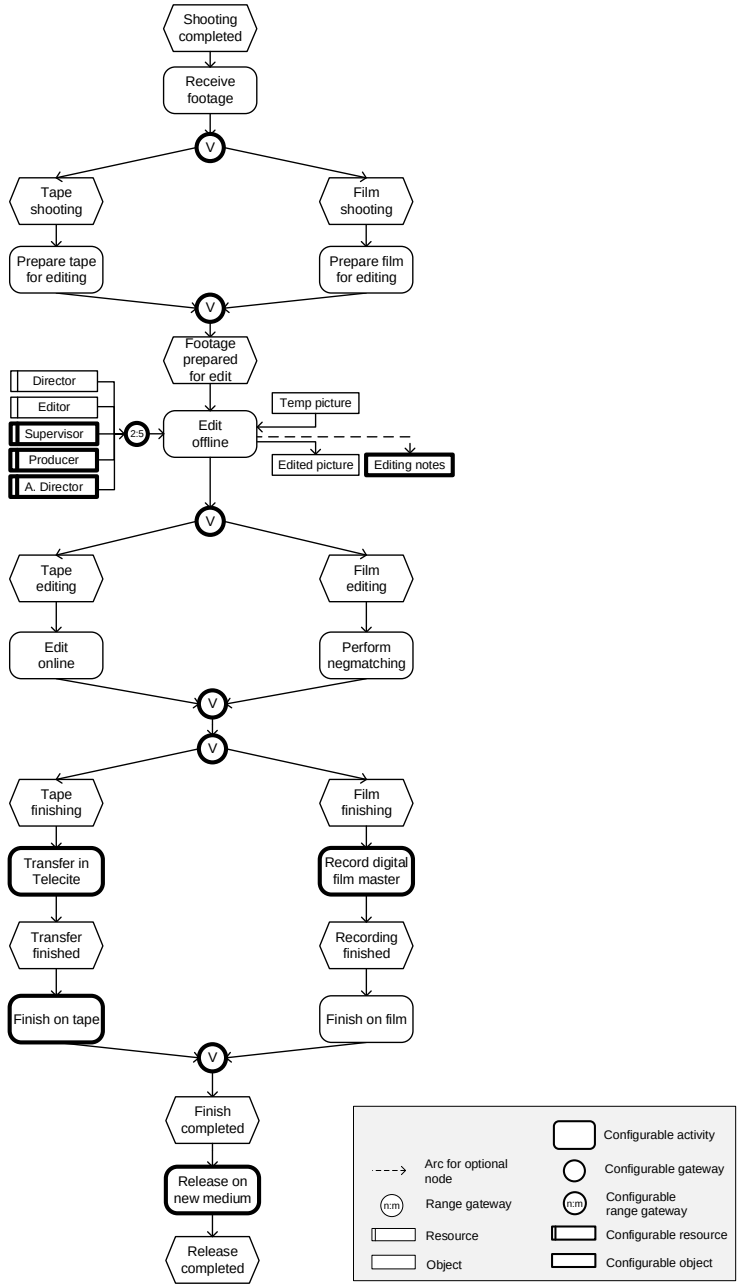

Fig. 4: The C-iEPC model representing all post-production variants.

as consumed, to indicate that it will be destroyed upon use by the activity. Moreover, resources and objects can be mandatory or optional, and can be connected to activities via logical gateways, called *range gateways*. Range gateways subsume the three logical types of OR, XOR and AND, but also allow any combination of the associated resources (objects), e.g. at least 2 and at most 5 resources. Range gateways can also be optional, in which case they indicate that all connected resources (objects) are optional.

For simplicity, Figure 4 only depicts the resources and objects associated with activity "Edit offline". This activity is performed by at least two resources, requires a Temp picture as input and produces an Edited picture as output. Editing notes are optional, since they might not be produced during the offline editing. In our example, three resources, one object and one range gateway have been marked as configurable with a thicker border. This fine-grained mechanism to allocate resources and objects to activities leads to different customization options. If a resource, object or range gateway is optional, it can be customized to *mandatory* so that is kept in the customized model, or switched *off*. If it is mandatory, it can only be switched off. Further, resources and objects can be specialized to a sub-type (e.g. a resource Producer can be specialized to an Executive Produced) according to a hierarchy model which complements the C-iEPC model (not shown in Figure 4). Configurable input objects that are consumed can be restricted to *used*, so that they are not destroyed by the activity after use.

C-iEPCs are a conceptual process modeling language – they do not provide any execution support. C-iEPCs are formally defined in [La Rosa et al. 2011]. The latter reference also defines an algorithm to derive an iEPC from a C-iEPC. If the C-iEPC is structurally correct, this algorithm preserves correctness when creating a customized model, by removing all nodes that are no longer connected to the initial and final events via a path, and by reconnecting the remaining nodes. Behavioral correctness is ensured via constraints inference (see Appendix D.1).

Abstraction and guidance during customization are achieved by means of a questionnaire, which captures domain properties and their values (see Appendix C.2), and is linked to the configurable nodes of a C-iEPC. C-iEPCs and associated questionnaire models are supported by the *Synergia*[3] and *Apromore*[4] toolsets. Using these toolsets, one can design C-iEPCs and questionnaire models, link these models, customize C-iEPCs via questionnaires and obtain the resulting customized models. The use of C-iEPCs has been validated via a case study in the film industry [La Rosa et al. 2011].

Table I summarizes the evaluation results for C-iEPCs. Each column indicates to what extent the approach in question covers each evaluation criterion defined in Section 3. We used a "+" on a green background to indicate a criterion that is fulfilled, a "−" on a red background to indicate a criterion that is not fulfilled and a "±" on an orange background to indicate partial fulfilment.

Scope					Customization Type		Supporting techniques				Extra-Functional		
Process Perspective			Process Type				Decision Support		Correctness Support				
Control-flow	Resources	Objects	Conceptual	Executable	Restriction	Extension	Abstraction	Guidance	Structural	Behavioral	Formalization	Implementation	Validation
+	+	+	+	-	+	-	+	+	+	+	+	+	+

Table I: Evaluation of C-iEPCs.

6.2. Configurable Workflows

The Configurable Workflows approach [van der Aalst et al. 2006; Gottschalk et al. 2007; Gottschalk et al. 2008] was first designed for conceptual models, and later applied to executable languages, with the aim to guarantee that the customized models

[3]See `www.processconfiguration.com`
[4]See `www.apromore.org`

can be executed. This led to the extension of several executable languages, such as SAP WebFlow, YAWL and BPEL. In this survey we focus on the extension to the YAWL language [Gottschalk et al. 2008], namely Configurable YAWL (C-YAWL), since this is the most significant one. The other extensions work in a similar way.

In YAWL, split and join gateways are graphically attached to activities: a join precedes an activity and models the activity's joining behavior; a split follows an activity and models its splitting behavior. C-YAWL extends YAWL with *ports* to identify configurable gateways. Configurable gateways are represented graphically with a thicker border, similar to C-iEPCs [van der Aalst et al. 2012]. A configurable split has an *outflow port* for each combination of subsequent flows that can be triggered after the activity completion, whilst a configurable join has an *inflow port* for each combination of sequence flows through which the activity can be triggered.

Figure 5.a depicts the post-production example in C-YAWL where for illustration purposes, we modeled the preparation and editing of tape and film as mutually exclusive activities. To illustrate the concept of port, let us consider the case of the first XOR-split, that of activity τ_1. This XOR-split is used to route the process flow according to the shooting media. This split can either give control to the top or to the bottom of its outgoing flows. Hence, given that the combination of outgoing flows is equal to the number of such flows for an XOR-split, this split has only two outflow ports, one to trigger the flow to event $0a$ (leading to the preparation of the film), the other to trigger the flow to event $0b$ (leading to the preparation of the tape).

Similarly, an XOR-join can be activated by each of its incoming flows, and so it has one inflow port for each incoming flow. This is the case of the XOR-join of activity "Edit offline" in our example. On the contrary, if the join (or split) is of type AND, it only has one inflow (or outflow) port. This is because an AND-join can only be activated by all its incoming flows (due to its synchronizing behavior), and similarly an AND-split gives control to all its outgoing flows simultaneously (due to its parallel behavior).

Let us now consider the OR-split of activity τ_2 in Figure 5.a. An OR-split has one outflow port for each combination of its outgoing flows, as it can give control to any combination of these flows. In our example, the OR-split is used to route control to activity "Record digital film master", or "Transfer in telecine" or both. Hence this OR-split has three outflow ports: one to trigger the flow to event $4b$, the other to trigger the flow to event $4a$ and the last to trigger both flows ($4a, 4b$).

The OR-join, on the other hand, only has one inflow port: this type of join is considered as an AND-join from a customization perspective, due to its synchronizing merge behavior. This is the case of the OR-join of activity "Release on new medium", which has a single inflow port to receive control from both its incoming flows.

Inflow ports have three customization options: *allowed*, *hidden* and *blocked*. An inflow port can be blocked to prevent the triggering of its activity, or hidden to skip the activity execution without blocking subsequent activities. So in C-YAWL both the semantics of hiding and removing an activity are supported. An inflow port that is neither blocked nor hidden, is allowed, i.e. it is kept as is in the customized model. An outflow port can only be blocked to prevent the triggering of the outgoing flows, or left allowed. For convenience, all the ports can be allowed or blocked by default. Activities can be customized via their inflow ports. Resources and objects cannot be customized.

Figure 5.a also shows a sample port customization for a project shot on tape, edited online and finished on film, overlaid on the C-YAWL model. Let us consider the first XOR-split. The only outflow port allowed by the example customization is the one that leads to activity "Prepare tape for editing". The inflow port from event $1a$ to the XOR-join of activity "Edit offline" is customized as blocked while the other inflow port for this join is allowed, to match the customization of the preceding XOR-split. Since the project is edited online, the outflow port of activity "Edit offline" triggering condition $2b$ is the only one to be allowed. In YAWL, an activity with a single incoming or outgoing

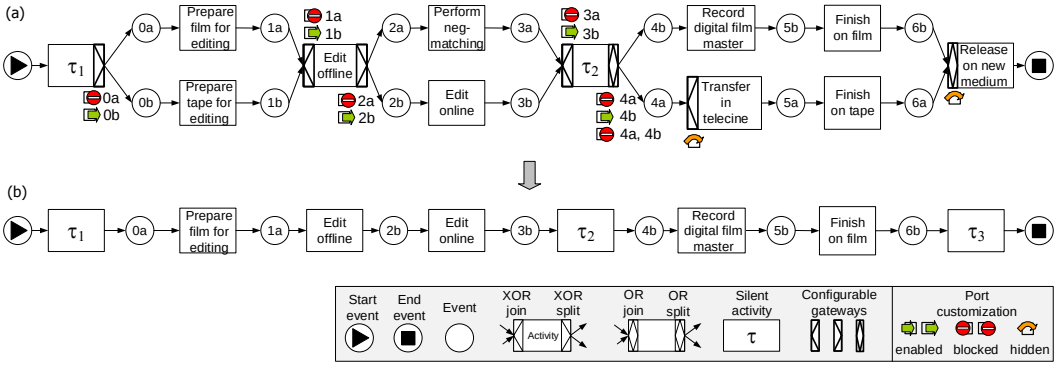

Fig. 5: (a) the post-production example in C-YAWL with a sample customization; (b) the customized model.

flow has an implicit XOR behavior. This behavior is shown graphically if the gateway must be made configurable, as in the case of the join of activity "Transfer in telecine", since we needed to hide this activity's inflow port.

The hiding and blocking operations can also be applied to other YAWL elements such as cancelation regions, composite activities and multi-instance activities.

Figure 5.b shows the YAWL model resulting from the example customization, after applying the transformation algorithm defined in [Gottschalk et al. 2008]. This algorithm removes all nodes that after customization are no longer on a path from the input to the output condition. In this way the structural correctness of the model is guaranteed. Moreover, potential conflicts in the data conditions of the outgoing arcs of (X)OR-splits are taken care of, in order for the resulting models to be fully executable. Two alternative techniques are available for ensuring the behavioral correctness of the customized models, one based on constraints inference and the other on partner synthesis, both described in Appendix D. Decision support is offered via the use of questionnaire models (see Appendix C.2).

This approach has been formalized [Gottschalk et al. 2008] and implemented in the *YAWL Editor*.[5] This tool allows one to create, customize and transform C-YAWL models into YAWL models, while support for customization via questionnaire models is offered by the *Synergia* toolset. The use of C-YAWL models with questionnaire models has been validated in the municipality domain [Gottschalk et al. 2009; Lönn et al. 2012], involving domain experts, as well as in software development processes for very small entities [Boucher et al. 2012].

Table II summarizes the evaluation results for Configurable Workflows.

Scope					Customization Type		Supporting techniques				Extra-Functional		
Process Perspective			Process Type				Decision Support		Correctness Support				
Control-flow	Resources	Objects	Conceptual	Executable	Restriction	Extension	Abstraction	Guidance	Structural	Behavioral	Formalization	Implementation	Validation
+	-	-	+	+	+	-	+	+	+	+	+	+	+

Table II: Evaluation of Configurable Workflows.

[5] See www.yawlfoundation.org

6.3. ADOM: Application-based Domain Modeling

In ADOM (Application-based Domain Modeling) [Reinhartz-Berger and Sturm 2007; Reinhartz-Berger et al. 2009; 2010] configurable nodes (activities, events and gateways), have a cardinality attribute of the form <min,max>. The cardinality specifies how many times a given node can be instantiated in the customized model. For example, an activity tagged with <0,1> is *optional*, and as such it can be dropped in the customized model; an activity tagged with <1,n> is *mandatory* and can be instantiated up to n times in the customized model; an activity tagged with <1,1> must be instantiated exactly once, i.e. it is kept as is in the customized model. The default cardinality <0,n> implies no constraints.

The cardinality assigned to gateways of type (X)OR indicates when the decision captured by the gateway should be made. A cardinality of <0,0> indicates that the gateway must not appear in the customized model. So at customization time one has to decide which outgoing branch(es) in the case of a split, or incoming branch(es) in the case of a join, to keep. If the cardinality is <0,1>, this decision can be deferred till run-time, i.e. the gateway is optional. An OR gateway can be restricted to become an AND or XOR, in the same way as in C-iEPCs and Configurable Workflows.

In ADOM, sequence flows can also be assigned a cardinality, unlike C-iEPCs and Configurable Workflows where sequence flows are not configurable. However, the customization of these flows is constrained by design by the customization of the configurable nodes, in order to avoid disconnections in the customized model. For example, a flow with cardinality <0,1> between two nodes with cardinality <0,1> cannot be dropped if the two nodes are kept, otherwise it would lead to a disconnection.

Figure 6.a shows the post-production example in EPCs with ADOM cardinality constraints. For example, event “Shooting completed”, activity “Receive footage” and the flow in-between are mandatory, so they can neither be removed nor instantiated more than once during customization. The OR-split and its matching OR-join are optional, and so are the nodes in-between. This is done to allow a choice between either of the two branches or both. All elements after activity “Edit offline” are mandatory but have a maximum cardinality greater than 1. By doing so, each of these elements can be instantiated multiple times to model the various options that exist for editing and finishing in post-production, though these options are not represented in the model.

In ADOM, commonalities between variants are thus captured by mandatory elements while variability is captured by optional elements and by those that can be instantiated multiple times. Since an ADOM model is meant to be used as a template, some parts can be left underspecified. During customization, each configurable element can be instantiated according to its cardinality constraint. Moreover, *application-specific* elements can be added anywhere. These elements only appear in the customized model, without any counterpart in the customizable model. Thus, ADOM supports customization by restriction (removing optional elements) and by extension (instantiating an element multiple times and adding application-specific elements).

In a customized model, each node that has been derived from a configurable node bears a *model classifier* (indicated between “<” and “>”), i.e. a reference to the originating node in the customizable model. If the label of the node needs to be changed, e.g. a more specific one is required, this can be added below the model classifier. Figure 6.b shows a possible customization of the post-production model, where application-specific elements are highlighted in gray. For example, the first two gateways have been obtained by restricting the type of the first two OR gateways to an AND. Event “Film editing” and activity “Perform negmatching” derive from event “Editing”, resp., activity “Edit”, and have been given each a new name. The second pair of AND gateways and the flow between activity “Transfer in telecine” and event “Transfer com-

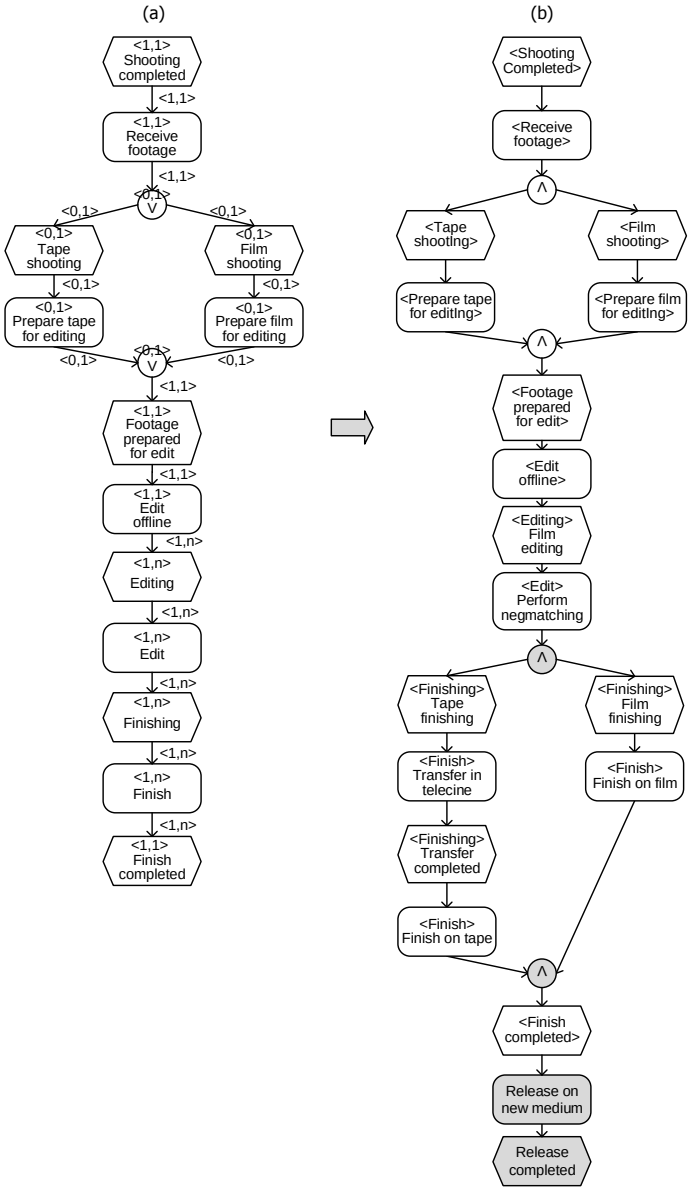

Fig. 6: (a) Post-production example in ADOM-EPC. (b) A customized model.

pleted", are application-specific elements added to allow multiple instantiations of event "Finishing" and function "Finish".

ADOM has been applied to the control flow of EPCs, UML Activity Diagrams (ADs) and BPMN at the conceptual level. For EPCs, specific rules have been defined to bind the customization of an event to that of an activity, in order to maintain the alternation between events and activities required by EPCs, though disconnected nodes cannot be avoided. Behavioral correctness of the customized models is not guaranteed.

Customization is performed directly on the model level. There is no means to specify which combinations of instantiations are unfeasible from a domain viewpoint, and the

addition of application-specific elements cannot be constrained. In [Reinhartz-Berger et al. 2009], the authors describe a validation technique for ADOM-BPMN. This technique checks a-posteriori that a customized model is compliant with its customizable model, but does not prevent the user from generating inconsistent or irrelevant customizations in the first place. As such, decision support is not offered.

A formalization of ADOM is provided in [Reinhartz-Berger et al. 2009] for BPMN and in [Reinhartz-Berger et al. 2010] for EPCs. The approach has not been implemented in a tool. A subset of ADOM-BPMN has been validated in the development of a process-driven service-oriented system, but without involving domain experts [Reinhartz-Berger et al. 2009].

Table III summarizes the evaluation results for ADOM.

Scope					Customization Type		Supporting techniques				Extra-Functional		
Process Perspective			Process Type				Decision Support		Correctness Support				
Control-flow	Resources	Objects	Conceptual	Executable	Restriction	Extension	Abstraction	Guidance	Structural	Behavioral	Formalization	Implementation	Validation
+	-	-	+	-	+	+	-	-	±	-	+	-	±

Table III: Evaluation of ADOM.

6.4. Recap

At the core, the approaches in this group allow different types of nodes in a customizable process model to be tagged as "configurable". A configurable node can be restricted at customization-time. Activities can be removed, gateways can be restricted (an OR gateway can be turned into an AND or XOR gateway), and their incident arcs can be blocked (dropped altogether) or made mandatory. The approaches differ in terms of the types of nodes that can be made configurable, the configuration options offered, as well as supporting techniques and extra-functional criteria.

C-iEPC is the only approach in this group that supports customization of data objects and resources; C-YAWL is the only one that provides execution support; and ADOM is the only one that supports customization by extension in addition to customization by restriction. C-iEPC and C-YAWL offer decision and correctness support, and fulfill the extra-functional criteria, whereas ADOM only partially does so.

7. GROUP 2: ELEMENT ANNOTATION

The three main approaches that fall in this group, Configurative Process Modeling, Superimposed Variants and aEPCs, rely on the graphical *annotation* of model elements with properties of the application domain. Model elements that can be annotated include control-flow nodes (activities, events and gateways), sequence flows, resources and objects. Those model elements that are annotated become variation points. Different approaches support different subsets of model elements. Domain properties are assigned to model elements via *domain conditions*, which are boolean expressions over domain properties (e.g. "low budget = true").

Customization is achieved by selecting domain properties. The selection may be done directly or be aided by a domain model such as a feature model or a product hierarchy. Based on this selection, the domain conditions are evaluated and those that are false lead to the corresponding model elements to be removed from the model. The required model transformation after the removal of the model elements is approach-specific.

7.1. Configurative Process Modeling

In configurative process modeling [Becker et al. 2004; Becker et al. 2007a; Delfmann et al. 2006; Delfmann et al. 2007; Becker et al. 2006; Becker et al. 2007c; Becker et al. 2007b] , customization is achieved by fading out model elements that are not relevant to a given business scenario. A set of domain properties, called *business characteristics*, are used to determine the available scenarios and later drive the customization.

In the case of post-production example, we can define a business characteristic "Shooting type" (ST) with values: "Tape" (T) or "Film" (F), and a characteristics "Budget Level" (BL) with values: "Low" (L), "Medium" (M) or "High" (H). The latter is a high-level characteristic since a choice on the budget typically affects multiple decisions in post-production. These characteristics are linked to the elements of a process model by means of domain conditions, which are logical expressions over the characteristics. The link is rendered graphically by encapsulating the model elements into a shaded box, to which the domain condition is attached (see Figure 7 for an example).

Process models are captured in eEPCs (extended EPCs), an extension of EPCs that incorporates resources and objects, similar to iEPCs. Business characteristics can be assigned to the following model elements: activities, events, resources and objects. Gateways cannot be directly configured. Rather, the approach expects the modeler to include all possible gateway variants in the customizable model.

Figure 7 shows a process model for post-production in eEPCs (the control flow is the same as that of the C-iEPC model of Figure 4). Here some elements have been associated with a logical expression referring to the project's budget. For instance, event "Film editing" and activity "Perform negmatching" are linked to the expression BL(H), which means that these elements are only suitable for a high budget project, due to the high costs involved in editing on film. On the other hand, activity "Edit online" is not associated with any condition, since it is suitable for any type of budget.

The customization of a process model to a specific scenario is done by marking those elements whose domain conditions evaluate to false as hidden. Then a transformation is performed to remove the hidden elements, including those gateways that become irrelevant, and reconnect the remaining nodes. For example, by customizing the example in Figure 7 for a low budget project, we obtain variant b in Figure 1. The transformation can fix simple structural issues, e.g. the removal of gateways which have one input and one output flow, as in the example, but cannot ensure the structural and behavioral correctness of the resulting models. For example, since both events and activities can be removed, the algorithm does not work in the presence of cycles, or when an activity between two events is removed. Structural issues that cannot be fixed are prompted to the modeler who has to manually correct them.

Business characteristics can also be applied to the meta-model layer (i.e. to the eEPC meta-model), to remove process modeling perspectives that are not relevant to a specific scenario. For example, one can hide all resources at once.

Customization is not carried out at the process model level, but via the evaluation of a set of business characteristics. However the approach does not offer guidance to users when assigning values to the characteristics.

The approach also supports a set of *generic adaptation mechanisms* that can be used to refine and extend a process model after customization, e.g. by adding new model fragments. The possible combinations of components can be restricted by interface definitions. However, the application of these mechanisms is left to the user without specific support, i.e., extension points are not specified in the customized process model. Hence, this approach only provides a weak form of customization by extension.

The approach builds upon eEPCs, which are a conceptual language. The approach has not been formalized. The model projection mechanism has been implemented as a prototype tool [Delfmann et al. 2006] that interacts with the ARIS platform. Business

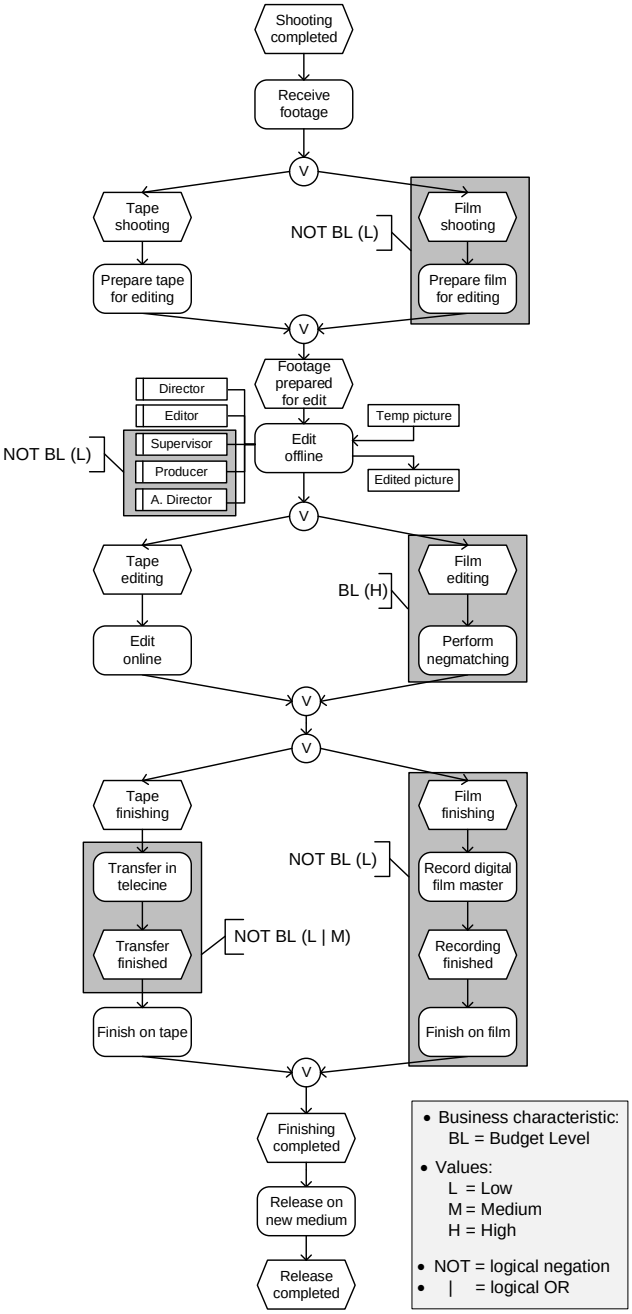

Fig. 7: A process model for post-production in eEPCs with logical terms for the budget.

characteristics can be defined and linked to elements of an eEPC designed in ARIS. Users select the desired characteristics and the initial eEPC is customized to remove irrelevant elements. Similar features are also available in the *[em]* tool.[6] This approach

[6] http://em.uni-muenster.de

has been applied in the fields of method engineering [Becker et al. 2007c] and change management [Becker et al. 2007b], and validated in the German public administration sector [Becker et al. 2006], though without domain expert involvement.

Table IV summarizes the evaluation results for Configurative Process Modeling.

Scope					Customization Type		Supporting techniques				Extra-Functional		
Process Perspective			Process Type				Decision Support		Correctness Support				
Control-flow	Resources	Objects	Conceptual	Executable	Restriction	Extension	Abstraction	Guidance	Structural	Behavioral	Formalization	Implementation	Validation
±	+	+	+	-	+	+	+	-	±	-	-	+	±

Table IV: Evaluation of Configurative Process Modeling.

7.2. Superimposed Variants

The idea of annotating model elements to capture variability is also investigated in [Czarnecki and Antkiewicz 2005; Czarnecki et al. 2005; Czarnecki and Pietroszek 2006]. In this approach, any control-flow element of UML ADs can be annotated using *presence conditions* and *meta-expressions*. A precedence condition determines if a model element is retained or removed. A meta-expression allows one to select the value of an attribute of a model element (e.g. an activity's label) from among a range of options. Customization is thus achieved by restriction only.

Both presence conditions and meta-expressions are captured by boolean formulae over the features and feature attributes of a feature model (see Appendix C.1), and are evaluated against a feature configuration. These formulae are represented in disjunctive normal form where the basic terms are features designated by means of UML stereotypes. For example, the stereotype <<`Tape` $\vee$ `Film`>> indicates the disjunction between features "Tape" and "Film". The assignment of stereotypes to modeling elements is done through rendering mechanisms such as labels, color schemes or icons.

Figure 8.a shows the finishing phase of the post-production process as an annotated UML AD. For simplicity, in this example we have only specified presence conditions. These annotations, rendered with a color and a number in the example, have been defined over the features of the feature model of Figure 15 (cf. Appendix C.1). This feature model captures the features (i.e. properties) of the post-production domain, such as type of finish and type of transfer. For example, activity "Transfer in telecine" is associated with the sub-feature "Telecine" of feature "Transfer" (annotated in blue with label "1"), while the two outgoing flows of the decision point are associated with the two sub-features of "Finish": "Tape", resp., "Film". All non-labeled elements (in black) are associated with the *always-true* formula. These represent the commonalities of the model and cannot be removed, e.g. the gateways and the end event in our example.

Customization is achieved by evaluating presence conditions and meta-expressions against a feature configuration. Model elements whose conditions evaluate to false are removed, while model attributes that are affected by meta-expressions are modified accordingly. No guidance is provided for the selection of the features to be kept.

Figure 8.b shows a possible customized model for the post-production example where only the activities "Record digital film master" and "Finish on film" have been kept. This model can be obtained via a transformation algorithm that applies *patches* to reconnect model elements that have been disconnected during customization, and *simplifications* to remove splits and joins that have been left with one incoming and one

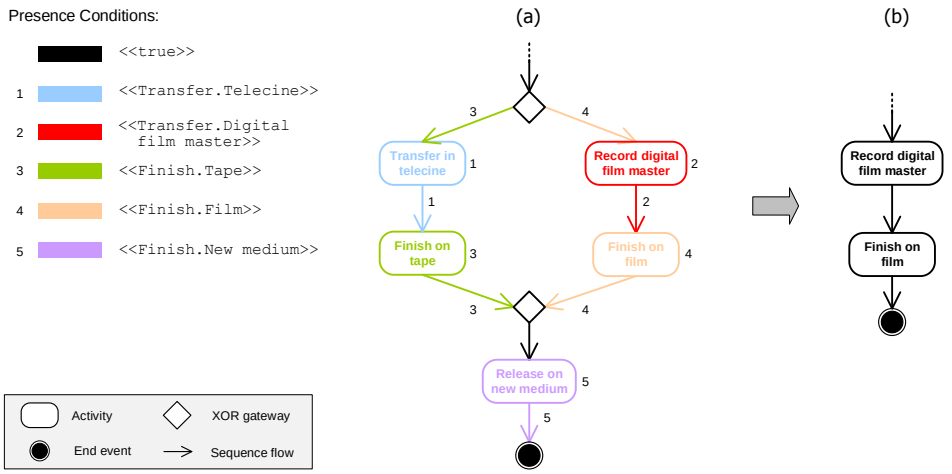

Fig. 8: (a) The post-production example in annotated UML ADs. (b) A customized model.

outgoing flow. Patches can only be applied to those nodes that have exactly one incoming and one outgoing flow, and an annotation error is raised otherwise. However, an automated verification procedure [Czarnecki and Pietroszek 2006] can be used to provide an a-priori guarantee that no structurally-incorrect customized model can be generated from a customization. Behavioral correctness is not dealt with, and no execution support is provided. The approach only supports customization of control-flow elements. Resources and objects cannot be customized.

The approach has been formalized and implemented in an Eclipse plugin[7] allowing users to configure UML ADs via so-called *cardinality-based feature models* [Czarnecki et al. 2005] and to check that the feature model and UML AD do not lead to structurally incorrect customized models. The approach has not been validated in practice.

Table V summarizes the evaluation results for Superimposed Variants.

Scope					Customization Type		Supporting techniques				Extra-Functional		
Process Perspective			Process Type				Decision Support		Correctness Support				
Control-flow	Resources	Objects	Conceptual	Executable	Restriction	Extension	Abstraction	Guidance	Structural	Behavioral	Formalization	Implementation	Validation
+	-	-	+	-	+	-	+	-	+	-	+	+	-

Table V: Evaluation of Superimposed Variants.

7.3. aEPCs: Aggregated EPCs

Aggregated EPCs (aEPCs) [Reijers et al. 2009] are an extension of EPCs to capture a family of process variants. Similar to Configurative Process Modeling and Super-

[7]See `http://gp.uwaterloo.ca/fmp2rsm`.

imposed Variants, the idea is to annotate certain model elements (in this case EPC activities and events) with domain properties, which are called *products* in aEPCs.

Figure 9.a shows an example aEPC where products associated with activities and events refer to the budget levels in post-production. For example, activity "Transfer in telecine" only occurs in high budget projects, while activity "Record digital film master" can also occur in medium budget projects. Accordingly, "High budget" and "Medium budget" are sub-products of a composite product "Budget" in post-production, i.e. they capture the values of a given domain property. Other possible products include the shooting formats, the picture cut methods and the finishing formats.

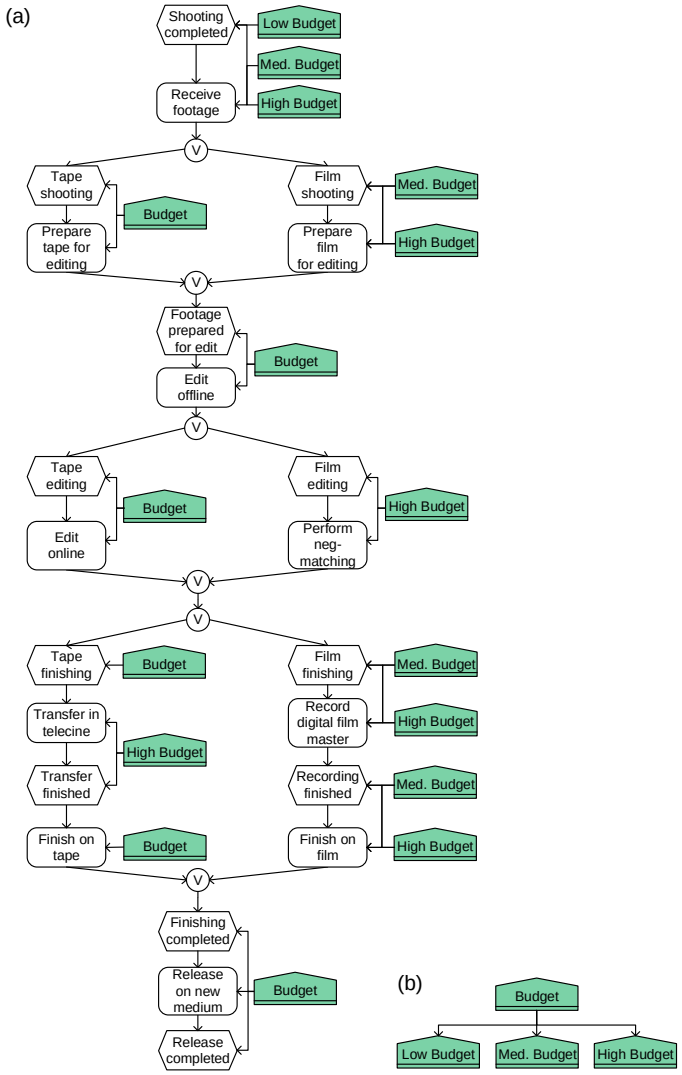

Fig. 9: (a) Post-production example in aEPC. (b) Associated product hierarchy for budget.

An activity or event may be associated with more than one product. In our example, activity "Record digital film master" is associated with two products ("High budget"

and “Medium budget”). In order to avoid cluttering the model with many product associations, an aEPC can be accompanied by one or more *product hierarchies* where the various products are organized hierarchically. A product hierarchy is a rooted tree where the leaves are products and all other nodes are composite products representing product generalizations. In this way a process model element can be associated with a composite product in place of a set of products. For example, Figure 9.b shows the product hierarchy for the budget. The composite product “Budget” in this hierarchy can be used when an element is present in all budget levels, e.g. activity “Receive footage”.

Instead of capturing implications among model elements or domain properties (e.g., “Edit online” can only be present if “Prepare tape for editing” is present) as in other approaches, in aEPCs all possible variants have to be resolved beforehand and mapped to a set of products (i.e. a composite product). Thus, while the use of composite products can in principle reduce the number of products associated with a given element, there may be a large number of composite products and this in turn may lead to cluttered aEPCs [Baier et al. 2010]. The choice of not modeling implications explicitly is motivated by the observation that in practice these logical expressions are difficult to conceive and interpret by domain experts. This was the result of testing C-EPCs (an ancestor of C-iEPCs) with domain experts of ING Investment Europe, with whom the aEPC approach was later validated [Reijers et al. 2009].

An aEPC is customized by choosing one or more products, and removing all elements that are not associated with the products chosen. Customization is restricted to activities and events, while gateways, objects and resources are not customizable.

This approach works at the conceptual level only since aEPCs are conceptual models. The approach is fully formalized including a transformation algorithm which removes the unneeded elements and cleans up the customized model, in order to keep it structurally correct. In fact, besides the requirements of an EPC, there are requirements on how products can be associated with elements appearing before or after a sequence of gateways. Behavioral correctness is not dealt with. The transformation algorithm has been implemented in a tool that can import EPCs from ARIS and extend them into aEPCs. An advantage of organizing products into hierarchies is that an aEPC can be customized by removing products from the associated product hierarchy. Thus, this approach achieves process abstraction, though guidance is not offered.

Table VI summarizes the evaluation results for aEPCs.

Scope					Customization Type		Supporting techniques				Extra-Functional		
Process Perspective			Process Type				Decision Support		Correctness Support				
Control-flow	Resources	Objects	Conceptual	Executable	Restriction	Extension	Abstraction	Guidance	Structural	Behavioral	Formalization	Implementation	Validation
±	-	-	+	-	+	-	+	-	+	-	+	+	+

Table VI: Evaluation of aEPCs.

7.4. Recap

Approaches in this group capture variability via annotations attached to elements of the customizable process model. These elements link an element in the customizable process model to an element in a domain model. Customization is performed by instantiating the domain model to capture a given set of requirements. Given an instance of

the domain model and the annotations in the customizable process model, a transformation algorithm is applied to derive a customized model. In Configurative Process Modeling, the domain model consists of business characteristics, in Superimposed Variants it takes the form of a feature model, and in aEPCs it consists of products.

The approaches differ in terms of the model elements that can be customized. All approaches support the customization of control-flow model elements, although Configurative Process Modeling is limited in its support for customization of gateways. On the other hand, Configurative Process Modeling is the only one that supports resources and objects. In all three approaches, customization is achieved by restriction, though Configurative Process Modeling also supports a weak form of extension. All three approaches provide abstraction support, since the customization is driven by domain concepts. However, none of them provides customization guidance. All three approaches ensure structural correctness (at least to some extent), but not behavioral correctness. They all target conceptual process models rather than executable ones.

8. GROUP 3: ACTIVITY SPECIALIZATION

The main approaches in this group, PESOA, BPFM and Feature Model Composition, rely on activity specialization to achieve process model customization. An abstract activity can be defined as a variation point by assigning one or more *variants* to it. A variant is a *specialization* of an abstract activity, i.e. one of its possible concrete refinements, e.g. activities "Prepare tape for editing" and "Prepare film for editing" are two specializations of "Prepare medium for editing". A special type of variation point is the *optional* activity: an abstract activity that can be specialized to an empty activity.

Variants can also be assigned to activity attributes such as objects and resources, which become variation points. Events and gateways cannot be customized. Accordingly, variability is graphically rendered by marking activities and their attributes as variation points and connecting variants to variation points via a specialization arc.

Customization is achieved by selecting one or more variants per variation point, while optional activities can be switched off. Customization can be done directly on the process model or via the use of a domain model, such as a feature model. The routing behavior to be used when selecting more than one variant for a variation point, as well as the transformation needed to clean up the model from all unused variants and to remove optional activities that have been switched off, are approach-specific.

8.1. PESOA: Process Family Engineering in Service-Oriented Applications

The idea of capturing variability in process models has been explored in the PESOA (Process Family Engineering in Service-Oriented Applications) project [Puhlmann et al. 2005; Schnieders and Puhlmann 2006; Schnieders 2006]. The aim of this project was not to provide a language for representing and customizing process models, but rather to improve the customization of process-oriented software systems, i.e. systems that are developed from the specification of process models. If the variability of a software system can be directly represented in a process model that describes the system's behavior, it is then possible to generate code stubs for the system from the customization of the process model itself. Since code generation is outside the scope of this paper, we only focus on the way the authors represent process variability.

According to PESOA, a customizable process model is a conceptual process model where certain activities have been marked with stereotypes to accommodate variability. Although stereotypes are an extensibility mechanism of UML, in this approach they are applied to both UML ADs and BPMN models. The activities of a process model where variability can occur are marked as variation points with the stereotype `<<VarPoint>>`. A variation point represents an abstract activity, such as "Prepare medium for editing", that needs to be specialized with a concrete variant (`<<Variant>>`) among a set of possible ones. For example, "Prepare medium for editing" can be spe-

cialized into “Prepare tape for editing” or “Prepare film for editing”, or both. One can also mark the default variant for a variation point with the stereotype <<Default>>. Figure 10.a shows the process model for post-production in BPMN, where some activities have been marked as variation points with their variants shown below the activity. For example, “Prepare tape for editing” is marked as the default variant of “Prepare medium for editing”, as this is the most common choice in post-production.

If the variants are exclusive, i.e. if only one variant can be assigned to a given variation point, the stereotype <<Abstract>> is used instead of <<VarPoint>>. In Figure 10.a we assume that the variants “Edit online” and “Perform negmatching” are exclusive, so the associated variation point “Cut picture” is marked with the tag <<Abstract>>. As a shortcut, when the variants are exclusive, the default specialization can be depicted directly on the variation point with the stereotype <<Alternative>>.

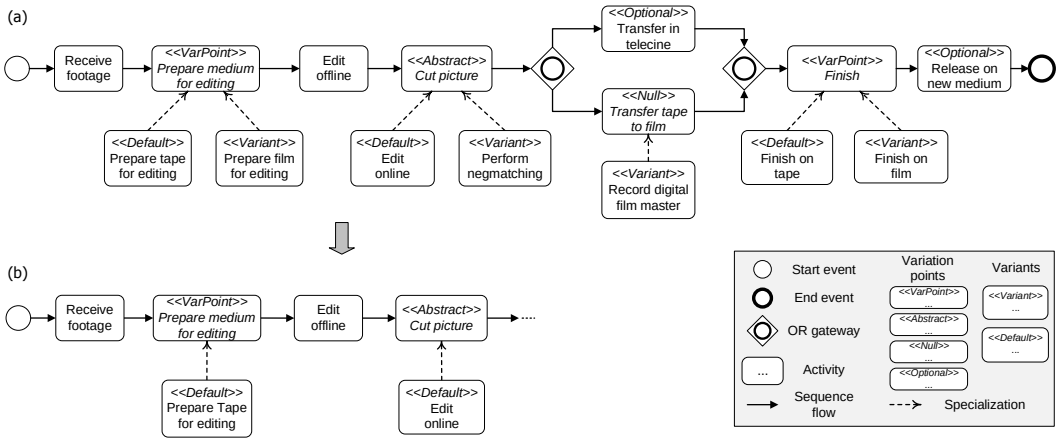

Fig. 10: (a) Post-production example in PESOA-BPMN. (b) A customized model.

A variation point marked with the stereotype <<Null>> indicates an optional activity. It can only be associated with one variant and its resolution into such variant is not compulsory, in which case the activity is switched off. This is the case of the variation point “Transfer tape to film” that may be specialized into the variant “Record digital film master”, or be completely dropped from the process model. A shortcut for a <<Null>> variation point and its variant is to depict the variant directly on the variation point, using the stereotype <<Optional>>, like task “Transfer in telecine”, which subsumes the variation point “Transfer film to tape”.

Stereotypes can be assigned to activities and to activity attributes related to objects (e.g. input and output data). Gateways, events and resources cannot be customized. During customization, each variation point is specialized into one or more variants depending to its type. Figure 10.b shows a fragment of the BPMN process model for post-production configured for a project shot on tape and edited online. The variants that are not required have been removed from the model. Extension mechanisms are not provided. Abstraction from the process modeling language is achieved by linking process variants with domain properties, captured as features in a feature model (see Appendix C.1). Each process variant is tagged with a feature, such that when a feature is disabled in a feature model configuration, the corresponding variant is removed from the process model. Domain constraints can be defined over feature values, thus restricting the possible combinations of variants in the process model. However, there is no guidance for the selection of a suitable set of features.

PESOA does not provide a transformation algorithm to derive customized models. The removal of certain variation points such as `<<Null>>` or `<<Optional>>`, as well as the customization of a variation point when multiple variants are selected, may lead to correctness issues which have to be fixed manually. A formalization is provided for selected concepts only [Puhlmann et al. 2005].

PESOA has been implemented as an Eclipse plugin. In this implementation, the customization of a process model is limited to removal of undesired variants. The approach has been validated in the hotel booking domain, in collaboration with ehotel and Delta Software Technology [Schnieders and Puhlmann 2006]. In this study, a set of BPMN process models were configured to drive the generation of Web applications in collaboration with domain experts.

Table VII summarizes the evaluation results for PESOA.

Scope					Customization Type		Supporting techniques				Extra-Functional		
Process Perspective			Process Type				Decision Support		Correctness Support				
Control-flow	Resources	Objects	Conceptual	Executable	Restriction	Extension	Abstraction	Guidance	Structural	Behavioral	Formalization	Implementation	Validation
±	-	+	+	-	+	-	+	-	-	-	±	±	+

Table VII: Evaluation of PESOA.

8.2. BPFM: Business Process Family Model

Business Process Family Model (BPFM) [Moon et al. 2008] is a two-level approach to capture customizable process models using an extended version of UML ADs. The first level deals with basic customization. At this level, an activity can be defined as *common* if it cannot be customized, or *optional* if it can be omitted during customization. The second level enable finer-grained customization by setting an activity as a *variation point* and assigning to it one or more specialized variants. Events, gateways, resources and objects cannot be customized.

A variant is a concrete activity; a variation point is an abstract activity of one of the following types: i) *boolean*, ii) *selection* or iii) *flow*. A boolean variation point can be specialized into exactly one variant. A selection requires at least one variant to be selected. In this case, the exact number of variants to be selected can be set with a cardinality (e.g. `1..2`). When selecting more than one variant, in BPFM one needs to specify the control-flow relation between the selected variants (called *flow pattern*). This can be a *sequence* (the selected variants are ordered sequentially), *parallel* (the selected variants are executed in parallel using an AND-split and an AND-join) or *decision* (they are made mutually exclusive using an XOR-split and an XOR-join). A flow variation point is assigned a *variants region*, i.e. a set of activities whose flow relations may be underspecified. At customization time, one needs to restrict the behavior by adding the required flows. A *flow pattern* can be specified for the flow variation point, in which case the activities in the variants region are organized according to the pattern, though the precise order needs to be decided by the user at customization time.

Further, the boundary of a variation point can be classified as either *closed* or *open*. A closed boundary restricts the choice of variants to those already identified whereas an open boundary allows the introduction of new variants during customization. Thus, in principle this approach supports both customization by restriction and extension. However, there is no support for plugging in new variants into a variation point.

Figure 11.a depicts the post-production example in BPFM. Here there are three activities marked with a variation point and one optional activity. Activities “Prepare medium for editing” and “Cut picture” are of type selection. They have been assigned two variants each. The first activity prescribes a parallel flow pattern while the second one a sequence flow pattern, each with the option of selecting at least one and at most two variants. Accordingly, Figure 11.b shows a customized model where the first variation point has been customized to the parallel execution of both its variants, whilst the second one to the sequence of its variants. Activity “Transfer & finish” is an open variation point of type flow with a decision pattern between the activities in the associated variants region. Accordingly, in Figure 11.b this variation point has been customized to a decision between two branches, each hosting two of the four activities present in the variants region. In case of an open flow variation point, the arrangement of activities inside the variants region within a given control-flow structure is entirely left to the user. Finally, in our example the optional activity “Release on new medium” has been dropped in the customized model.

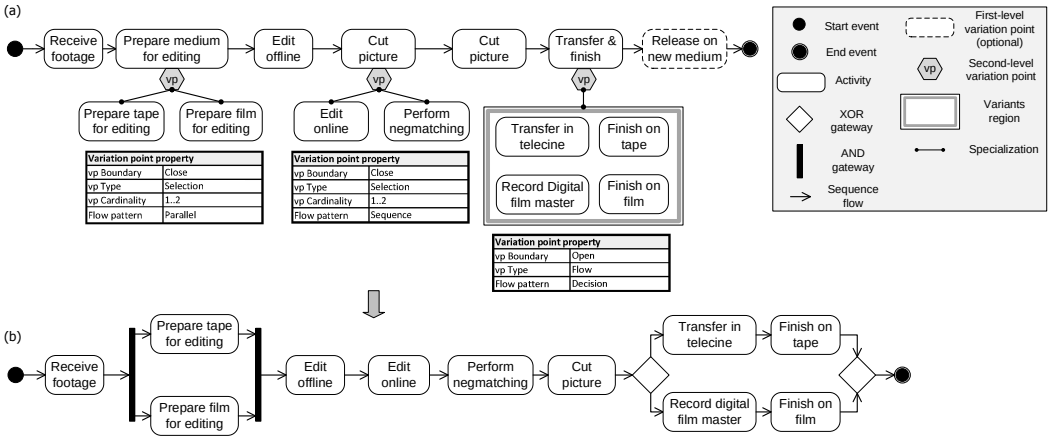

Fig. 11: (a) Post-production example in BPFM. (b) A customized model.

In BPFM it is also possible to define dependencies (called *dependency constraints*) between variation points, between variants or between variation points and variants. If for example, a variant is chosen for a given variation point, this can restrict the choice of the variants for another variation point.

A tool implementing this approach is available as an Eclipse plugin. The tool can prune a customized process model by removing the unused variants, but does not offer a complete transformation algorithm for embedding the selected variants into the process model. The approach has not been formalized nor validated in practice. It does not provide any correctness, abstraction or decision support.

Table VIII summarizes the evaluation results for BPFM.

8.3. Feature Model Composition

In Feature Model Composition [Acher et al. 2010b], a process model (called *workflow*) is defined as a collection of activities (called *services*). Activities are implicitly related via data dependencies. Specifically, each activity has a collection of attributes called *dataports*. A dataport captures an input or an output data object of the activity. If an input dataport of an activity refers to the same object as an output dataport of another activity, there exists an implicit data dependency between these two activities.

Scope					Customization Type		Supporting techniques				Extra-Functional		
Process Perspective			Process Type				Decision Support		Correctness Support				
Control-flow	Resources	Objects	Conceptual	Executable	Restriction	Extension	Abstraction	Guidance	Structural	Behavioral	Formalization	Implementation	Validation
±	-	-	+	-	+	+	-	-	-	-	-	±	-

Table VIII: Evaluation of BPFM.

In order to capture variability, an activity is allowed to have any number of variation points (called *concerns*). A concern refers to any activity attribute. Examples of attributes are dataports, functional interfaces, activity behavior and other low-level implementation aspects. Each concern is modeled as a separate feature model, which captures the variants that exist for a concern, and their relations. Customization of concerns is achieved by deselecting features from the respective feature models.

Figure 12 shows a customizable process model in the Feature Model Composition approach using our running example. A feature model has been defined for the concern "Shooting medium" and mapped to the output dataport of "Prepare medium for editing", in order to capture the fact that this activity can have a film, a tape or both media as output. Similarly, the same feature model has been associated with the input dataport of the subsequent activity "Edit offline". Further, the concern "Cut" with variants "Online" and "Negmatching" has been associated with the functional interface of activity "Perform cut", to indicate that the type of this activity can also be configured.

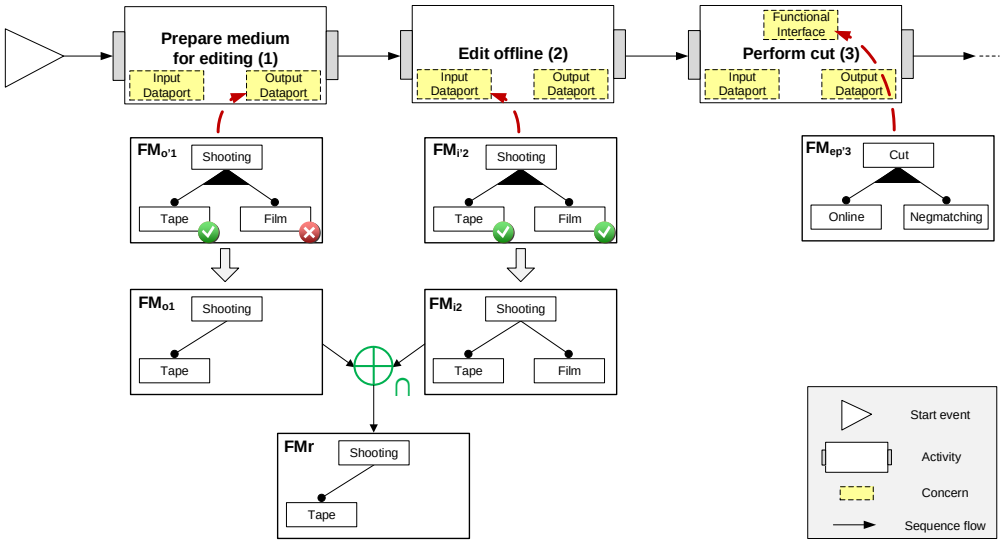

Fig. 12: Post-production example using Feature Model Composition and a possible customization.

A concern of one activity may be incompatible with that of a subsequent activity and thus a consistency check is needed when customizing a model. This check is performed by analyzing input and output dataports based on dependency rules. Specifically, the

feature models of the relevant concerns are checked for mutual consistency and then a merged diagram is created by intersecting the various feature models. In this way, the consistency of two connected activities is ensured. The merge operator is used to compose feature models that refer to the same activity dimension. Its syntax and semantics are defined in [Acher et al. 2010a], while the syntax of a customizable process model is defined in [Acher et al. 2010b]. While inconsistencies in data dependencies that may arise during customization are addressed by this approach, the process modeling language adopted is abstract, and not actually executable.

When producing a customized model, it is necessary to add the control-flow dependencies based on the implicit dependencies of the various activity attributes, such as data dependencies. Three types of control-flow dependencies are possible: sequential, concurrent (AND behavior) and conditional (XOR behavior). The dependency rules for consistency checks between two activities (cf. Figure 12) are not sufficient when there is a sequential, concurrent or conditional ordering of more than two activities. This is addressed via an extended set of dependency rules that ensures the consistency of the activities in a process model.

As shown in the example, this approach can be used to customize business objects and other activity attributes, such as the associated resources. However, concerns are internal to each activity. As such, the control flow cannot be configured. This is the only evaluated approach that suffers from this limitation.

In this approach, feature models do not provide abstraction for the customization of concerns, since they refer to low-level aspects such as different dataports related to a software service. Moreover, one has to configure one feature model per concern. There is no overarching feature model to customize the process model using properties of the application domain, like e.g. in PESOA. Similarly, no guidance support is offered.

Since the control flow cannot be configured, and data dependencies are preserved during customization, the approach guarantees that the customized models are both structurally and behaviorally correct.

An implementation is described on-line,[8] though the tool cannot be downloaded and the authors have not replied to our request for assistance with their tool. the Feature Model Composition approach is motivated by the need of customizing medical imaging grid services, though it has not been validated in practice.

Table IX summarizes the evaluation results for Feature Model Composition.

Scope					Customization Type		Supporting techniques				Extra-Functional		
Process Perspective			Process Type				Decision Support		Correctness Support				
Control-flow	Resources	Objects	Conceptual	Executable	Restriction	Extension	Abstraction	Guidance	Structural	Behavioral	Formalization	Implementation	Validation
-	+	+	+	±	+	-	-	-	+	+	+	+	-

Table IX: Evaluation of Feature Model Composition.

8.4. Recap

A distinctive characteristic of approaches in this group is they only allow customization of individual activities and not of other control-flow elements (events and gateways).

[8] See http://modalis.polytech.unice.fr/softwares/manvarwor.

Feature Model Composition does not even support the customization of an activity itself, but rather focuses on the customization of an activity's inputs and outputs. In other words, every activity in the customizable process model will appear in every customized model thereof. The customized models only differ in terms of the involved resources and data objects. Control-flow relations between activities have to be specified over the customized process model based on the data dependencies between activities.

Approaches in this group focus on conceptual process models. Since specialization is a form of behavior restriction, the approaches support customization by restriction. BPFM also supports a weak form of extension. PESOA provides abstraction support, but none of the approaches provides guidance. The approaches in this group do not ensure structural nor behavioral correctness, except for Feature Model Composition, which trivially achieves correctness support since it does not capture control-flow dependencies between activities.

9. GROUP 4: FRAGMENT CUSTOMIZATION

Approaches in this group are based on the application of *change operations* to restrict or extend the customizable process model. Two atomic change operations can be used to customize the control flow: *delete*, to remove a fragment from the model, and *insert*, to add a fragment into the model. More complex operations such as *move* or *replace* can be provided by combining delete and insert. The fragment to be deleted or inserted must be single-entry single-exit. Accordingly, each operation requires two sequence flows of the process model to delimit the portion of the base model to be deleted or inserted (the two flows may coincide). These variation points (called *adjustment points*) may be explicitly represented by marking selected flows of the model, otherwise each flow is assumed to be an adjustment point. The required model transformation after the application of the change operations is approach-specific. A third change operation, *modify*, is used to modify the resources or objects associated with an activity, e.g. replacing a resource with another or assigning a new resource to an activity.

Operations can be organized in an *operation sequence*, so that multiple operations can be performed in a given order on the customizable process model. Moreover, these sequences can be associated with *domain conditions*, i.e. predicates over domain properties, to determine when the sequence of operations in question should be applied.

This group counts two main approaches: Provop and Template and Rules.

9.1. Provop: Process variants by options

In Provop [Hallerbach et al. 2008; 2009a; 2009b; 2010], customization is achieved by applying change operations to a *base model* marked with adjustment points. The base model can be a standard process (e.g. a reference model for a particular domain), the most frequently used process variant, a generic model, the superset of all variants or their intersection. For example, in Figure 13 we identified variant a from the set of post-production variants in Figure 1 as the base model, since this is one of the simplest variants for post-production, and defined eight adjustment points on this model.

Besides the two atomic change operations for the control flow (DELETE and INSERT), and the MODIFY operation to customize objects and resources, Provop supports a fourth operation, namely MOVE, to relocate a fragment delimited by two adjustment points in the base model to another part of the model delimited by two different adjustment points. Operation sequences are called *options* in Provop.

For example, the DELETE operation in Option 1 of our example will delete the content between adjustment points "w" and "z". As a shorthand notation, it is possible to delete a single node simply by providing its identifier. A fragment is inserted in parallel to the portion of the base model delimited by two given adjustment points, if this portion contains some node. For example, in the case of the first INSERT of Option 4, an AND-split and an AND-join are used to link the fragment to the adjustment points.

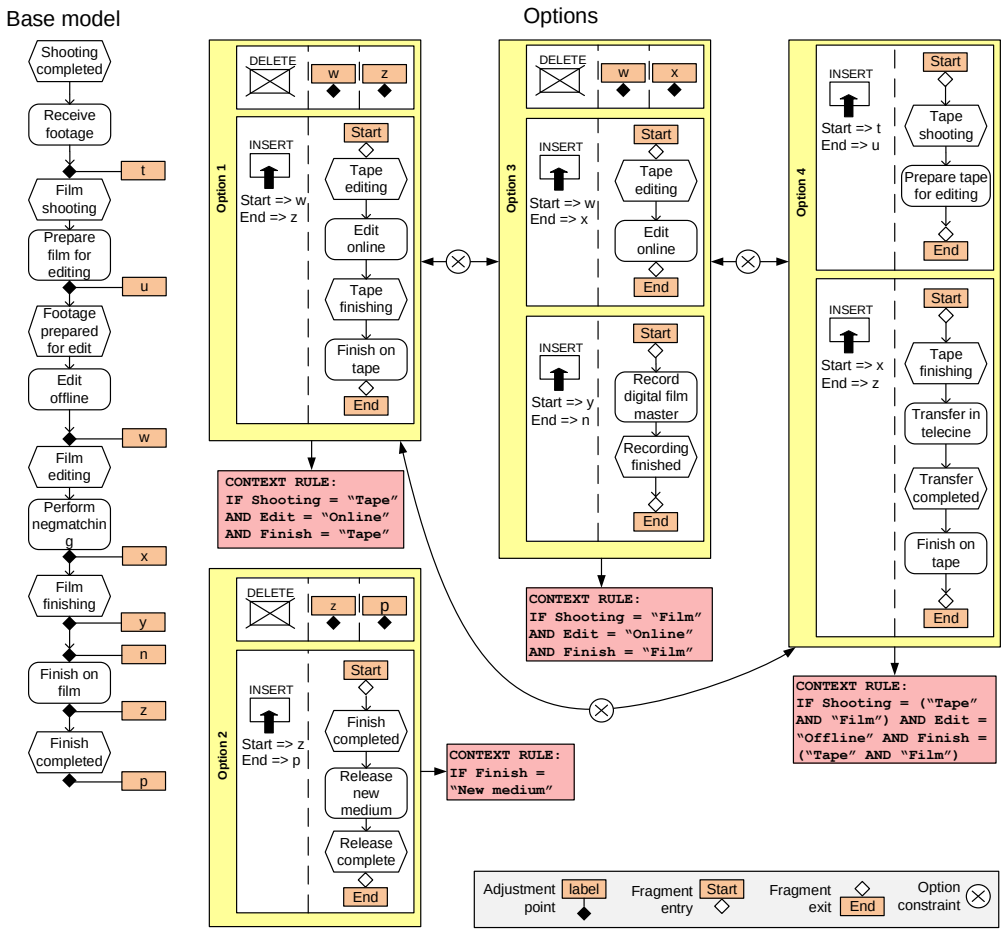

Fig. 13: Post-production example in Provop.

If the portion contains a sequence flow only, or is empty (e.g. as a result of a previous DELETE), the fragment is inserted in place of the flow or between the two adjustment points, respectively. An example of this is the second INSERT of Option 2, where the sequence "Record digital film master"–"Recording completed" is inserted in place of the flow between "y" and "n".

Since adjustment points can only be defined on the control flow, in Provop it is not possible to represet variability in the resource and object perspectives.

We organized the change operations in our example in four options. The application of Options 1 and 2 on the base model yields variant b of post-production, Option 3 yields variant c while Option 4 yields variant d (cf. Figure 1). The use of certain combinations of options can be restricted by defining *option constraints*, such as mutual exclusion, implication and n-out-of-m choices. For example, Options 1 and 3 of our example are set as mutually exclusive, since Option 1 removes the adjustment point "x" required by Option 3. The rationale behind the use of these constraints is to avoid creating situations that may prevent the application of an option or that may introduce errors in the resulting variants.

A five-step method can be used to drive the customization of process models via properties of the application domain [Hallerbach et al. 2009a; 2010]. In Step 1, the user

determines all the possible *contexts* in the application domain. A context is a domain property, represented as a variable, such as "budget", with all its possible values, e.g. "high", "medium" and "low". One can also specify domain constraints (called *context constraints*) in the form of boolean expressions to limit the interplay among contexts, e.g. "budget = low ⇒ finish = tape". Each option is then assigned a domain condition (*context rule*), in the form of a boolean expression over the values of context variables, to limit the applicability of that option to a particular business scenario. For instance, Option 1 can only be applied if shooting and finish are done on tape, and editing is done online. In Step 2, for each context the set of relevant options is automatically determined. In Step 3, the consistency of the retrieved options for each context is checked against the option constraints. If inconsistencies are found, these are prompted to the user, e.g. an option constraint may contradict a context constraint. In Step 4 all valid sets of options are applied to the base model for each context, and the resulting variant is checked for correctness in Step 5. Those models that are incorrect are discarded. Contexts and context rules offer abstraction for the customization of the base model, though guidance in the selection of the various options is not provided.

Provop can be applied to any modeling language with the only structural restriction that the fragments to be customized must be single-entry single-exit. The base model is not required to be correct. This however cannot guarantee the correctness of the customized model a priori. For example, the model may have disconnections or splits and joins of different type in a given single-entry single exit fragment, leading to behavioral anomalies. Instead, correctness is checked a posteriori (in Step 5) using existing correctness-checking techniques. In fact, the number of valid combinations of context variables into contexts may be very large, making an a-apriori check of all derivable customized models unfeasible in such cases.

The approach only addresses customization of conceptual process models. The basic concepts are formalized, though the semantics of the change operations is not specified. Provop has been implemented on top of ARIS [Hallerbach et al. 2010]. This tool allows users to define change operations and organize them in options, and to apply them to BPMN models enhanced with adjustment points, in order to derive customized models. The tool is not publicly available. Provop's design requirements have been derived from various case studies in the automotive and healthcare industries [Hallerbach et al. 2010], and Provop models have been created by the authors in these domains. However these models have not been validated with domain experts.

Table X summarizes the evaluation results for Provop.

Scope					Customization Type		Supporting techniques				Extra-Functional		
Process Perspective			Process Type				Decision Support		Correctness Support				
Control-flow	Resources	Objects	Conceptual	Executable	Restriction	Extension	Abstraction	Guidance	Structural	Behavioral	Formalization	Implementation	Validation
+	±	±	+	-	+	+	+	-	-	-	±	+	±

Table X: Evaluation of Provop.

9.2. Template and Rules

Template and Rules [Kumar and Yao 2009; 2012] captures the variability of a process family by processing a set of *business rules* associated with a *process template*. The process template is the customizable process model: a simple, block-structured process

model which should be chosen in order to have the shortest structural distance from all process variants of the family. The rules are sequences of change operations used to customize the template by restricting or extending its behavior. Change operations affect the control flow (by deleting, inserting, moving or replacing a fragment), the resources (by assigning a resource to an activity or changing the value of a resource property), and the data objects (by assigning a value to a data attribute, or changing the value of an activity's input/output data). The operations on resources and objects are "modify" operations, while those on the control flow are (a combination of) delete and insert. Unlike Provop, gateways cannot be directly customized, and adjustment points are not explicitly represented, meaning that change operations can virtually be applied to any process model fragment. As a result, though, in Template and Rules variability is not graphically represented in any process model perspective, and can only be inferred from the rules accompanying the template.

Rules are assigned domain conditions (e.g. "process.budget = high"), which if satisfied, allow the corresponding change operations to be applied to the process template. The use of such conditions provides abstraction from the customizable process model, though there is no guidance support.

Figure 14 shows the application of this approach to our running example, using the BPMN language. Here the template describes a simple variant for editing and finishing on tape and releasing on new medium. This template is accompanied by three rules (*R1*, *R2* and *R3*) embracing control-flow and resource aspects. For example, *R1* is used to customize the template for a high budget production process. Accordingly, we need to insert the activities required for editing and finishing also on film, such as "Prepare film for editing", to be inserted in parallel to "Prepare tape for editing", "Transfer in telecine", which goes before "Finish on tape" and so on (where *insert*(t_1, P, t_2) in a rule indicates to insert activity t_1 in parallel to t_2 while *insert*(t_1, S_b, t_2) indicates to insert t_1 before t_2). *R3* is an example of a rule to configure resource aspects: if the budget is high, multiple resources (e.g., "Director", "Editor", "Supervisor") will perform activity "Edit offline". Predicate *role*(t, r) indicates that resource r is assigned to activity t.

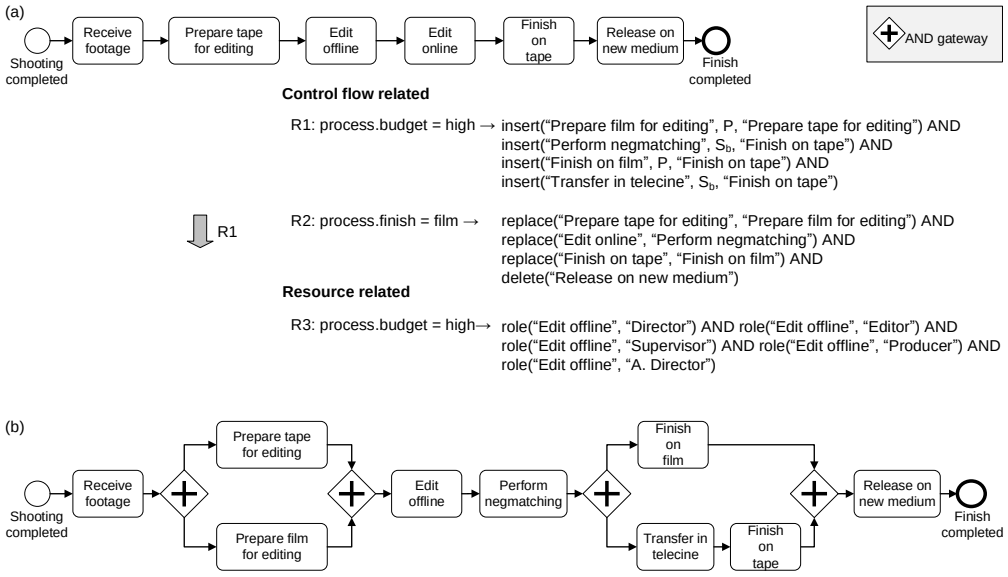

Fig. 14: (a) Post-production example in Template & Rules. (b) Customized model.

Change operations are applied to a tree representation of the template and, similar to Provop, only affect single-entry single-exit fragments of the template. Moreover, the application of each change operation triggers some cleaning operations to avoid disconnected model elements and remove trivial gateways and sequence flows. For example, after deleting activity "Release on new medium" from the template in Figure 14 via the application of rule *R2*, activity "Finish on tape" and event "Finish completed" will be reconnected. Similarly, if there remains one branch only between two AND gateways, the two gateways will be removed altogether. Thus, change operations cannot cause any structural nor behavioral issues in the process template.

Change operations are described in detail in terms of changes to the tree representation of the template and an algorithm is provided to customize the tree. However, a formalization of all notions is missing. Also, an algorithm to transform a process model into a tree representation and vice-versa is missing, while rule conflict resolution is only exemplified by a matrix that disallows certain combinations of rules.

The approach has been implemented using BPEL as the base language, though the tool is not publicly available. Given a template and a set of rules, the tool uses the Drools-expert rule engine to check for conflicts between the available rules. If conflicts exist (e.g. one rule deletes an activity another rule is trying to insert), the user is notified to either resolve them or assign a priority to each rule. The applicability of each rule is checked (e.g. it is not possible to delete an inexistent node) and errors are triggered for those rules that are not applicable. Finally, a customized process model is obtained from the template by only using those rules that are non-conflicting and applicable. This model is checked for data-flow inconsistencies, e.g. a task whose data input is no longer available, in order to guarantee the executability of the customized model. This check is done a-posteriori, as a result of which a customized model may be unfeasible altogether. The approach has not been validated in practice.

Table XI summarizes the evaluation results for Template and Rules.

Scope					Customization Type		Supporting techniques				Extra-Functional		
Process Perspective			Process Type				Decision Support		Correctness Support				
Control-flow	Resources	Objects	Conceptual	Executable	Restriction	Extension	Abstraction	Guidance	Structural	Behavioral	Formalization	Implementation	Validation
±	±	±	+	+	+	+	+	-	+	+	±	+	-

Table XI: Evaluation of Template and Rules.

9.3. Recap

The approaches in this group capture variability by means of (sequences of) change operations applied to the customizable process model. These change operations can add, delete or modify. Hence, the approaches in this group naturally support both customization by restriction and by extension. Unlike Provop, Templates and Rules supports both conceptual and executable models and supports abstraction as well as structural and behavioral correctness (at the expense on structural restrictions on the types of fragments that can be deleted or added). The combination of these characteristics makes Templates and Rules stand out in terms of its comprehensive coverage of the evaluation criteria.

10. DISCUSSION

This section compares the surveyed approaches (including the subsumed ones) in terms of the criteria introduced in Section 3. This comparison is followed by a discussion on the research questions introduced in Section 1.

10.1. Comparative analysis

The comparative analysis of approaches is summarized in Table XII. The first column lists the eleven main approaches and the twelve subsumed approaches. The next three columns indicate the year of the primary publication, the total number of citations (including all papers related to a given approach), and the modeling language(s) employed by the approach. The remaining columns indicate the coverage of each criterion.

	Approach	Year of primary publication	Total citations	Process modeling language	Scope					Customization Type		Supporting techniques				Extra-Functional		
					Process Perspective			Process Type				Decision Support		Correctness Support				
					Control-flow	Resources	Objects	Conceptual	Executable	Restriction	Extension	Abstraction	Guidance	Structural	Behavioral	Formalization	Implementation	Validation
Main	C-iEPCs	2003	1,313	C-iEPCs	+	+	+	+	-	+	-	+	+	+	+	+	+	+
	Configurative Proc. Modeling	2004	278	eEPCs	±	+	+	+	-	+	+	+	-	±	-	-	+	±
	PESOA	2005	226	BPMN, UML ADs	±	-	+	+	-	+	-	+	-	-	-	±	±	+
	Superimposed Variants	2005	1,287	UML ADs	+	-	-	+	-	+	-	+	-	+	-	+	+	-
	Configurable Workflows	2006	772	C-YAWL, C-SAP, C-BPEL	+	-	-	+	+	+	-	+	+	+	+	+	+	+
	ADOM	2007	125	UML ADs, EPCs, BPMN	+	-	-	+	-	+	+	-	-	±	-	+	-	±
	BPFM	2008	22	UML ADs	±	-	-	+	-	+	+	-	-	-	-	-	±	-
	Provop	2008	577	Any	+	±	±	+	-	+	+	+	-	-	-	±	+	±
	aEPCs	2009	90	aEPCs	±	-	-	+	-	+	-	+	-	+	-	+	+	+
	Template and Rules	2009	52	Block-structured BPEL	±	±	±	+	+	+	+	+	-	+	+	±	+	-
	Feature Model Composition	2010	29	Any	-	+	+	+	±	+	-	-	-	+	+	+	+	-
Subsumed	KobrA	2000	297	UML ADs	+	-	-	+	-	+	-	+	-	-	-	-	+	±
	Ciuksys & Caplinskas	2006	15	UML ADs	±	-	-	+	-	+	-	+	-	-	-	-	-	-
	Korherr & List	2007	33	UML ADs	+	-	-	+	-	+	-	+	-	-	-	-	-	-
	Razavian & Khosravi	2008	55	UML ADs	±	-	+	+	-	+	-	-	-	-	-	-	-	-
	Kulkarni & Barat	2010	15	BPMN	±	-	-	+	-	+	-	-	-	-	-	±	-	-
	Ripon et al.	2010	10	UML ADs	±	-	-	+	-	+	-	+	-	-	-	-	-	-
	Santos et al.	2010	22	BPMN	+	+	-	+	-	+	+	+	-	-	-	-	-	-
	CoSeNet	2011	34	CoSeNets	+	-	-	+	±	+	-	-	-	+	+	±	+	±
	Machado et al.	2011	16	BPMN	+	-	-	+	±	+	+	+	-	-	-	+	±	-
	Nguyen et al.	2011	27	BPMN	±	-	+	+	-	+	-	+	-	-	-	-	+	-
	vBPMN	2011	36	Block-structured BPMN	+	-	-	+	±	-	+	-	-	+	+	+	+	±
	Gröner et al.	2013	22	Block-structured BPMN	±	-	-	+	-	+	-	+	±	-	-	±	+	±

Table XII: Evaluation results at a glance, ordered by year of primary publication.

Regarding the modeling scope, all approaches (except Feature Model Composition) provide customization mechanisms along the control-flow perspective, but only a handful support the customization of resources and objects. Approaches based on BPMN and UML ADs do not support the customization of resources, except for Santos et al. This is probably because these two languages provide limited support for capturing resources, beyond the ability to associate a lane or a pool with each activity in the process. Accordingly, it is mainly in the context of EPCs or other languages that the question of customization of resources is posed. Customization of objects, on the other hand, is available in different languages, but only one (Templates and Rules) addresses customization of data objects in the context of executable process models.

In a similar vein, most approaches are based on conceptual modeling languages (UML ADs, EPCs, BPMN), and are hence focused on the customization of conceptual rather than executable process models. BPMN version 2.0 supports the specification of executable processes, but no customization approach so far covers the executable

features of BPMN (e.g. customization of data variables). Configurable Workflows and Template and Rules are the only approaches that fully support customization of executable models (in YAWL, BPEL and SAP WebFlow), down to the level of producing models that can be deployed in a BPMS. One can hypothesize that the observed emphasis on conceptual process modeling stems from the fact that variability in executable process models is usually tackled via run-time customization [Reichert and Weber 2012] rather than design-time customization (cf. Section 2).

All but one approach (vBPMN) support customization by restriction, while only a minority of approaches support customization by extension (eight out of 23). There appears to be a tradeoff between supporting customization by extension and preserving correctness. Indeed, approaches that support customization by extension do not support correctness, except for Template and Rules and vBPMN, which support correctness at the expense of imposing constraints on the structure of the customizable model and allowed extensions, namely that they both must be block-structured.[9] This observation highlights the fact that in order to reconcile customization by extension and correctness support, it is necessary to constrain the allowed extensions and the places in the customizable process model where these extensions can be inserted.

CoSeNet also achieves correctness support at the expense of structural constraints on the customizable process models (block-structured). On the other hand, C-iEPCs and Configurable Workflows achieve both structural and behavioral correctness without imposing structural constraints. This is achieved via incremental checks that detect combinations of customization options leading to incorrect models. However these approaches only allow customization by restriction, in line with the observation above.

The majority of approaches support customization based on domain models (i.e. abstraction), which may take the form of predicates over domain properties (as in Configurative Process Modeling and Provop), feature models, questionnaire models or decision tables as discussed in Appendix C. On the other hand, only two approaches (C-iEPCs and Configurable Workflows) provide step-by-step guidance to make customization decisions while avoiding inconsistent or irrelevant decisions to be taken. The approach by Gröner et al. does not provide step-by-step guidance, but prevents inconsistencies between decisions made during customization.

It is positive that the majority of approaches have tool implementations, at least partially, and about half of the approaches are fully or partly formalized. Also, about half of the approaches have been validated at least partially using real-life scenarios, although in many cases the validation has not involved domain experts. Overall, these observations highlight the relative maturity of the field.

C-iEPCs and Templates and Rules come close to supporting all the criteria. C-iEPCs focus on customization by restriction in conceptual process models. Templates and Rules covers both conceptual and executable models, but leaves aside the issue of providing customization guidance. These approaches demonstrate that the identified criteria are rather orthogonal, meaning that it is possible to support all of them. The only partial tradeoff is the one between supporting customization by extension and supporting correctness preservation. This tradeoff however is not necessarily unsurmountable. One can conceive approaches that achieve correctness preservation while supporting customization by extension, by setting boundaries on the way the customizable process model can be extended, e.g. only certain pre-defined templates can be employed and these templates are defined in a way that behavioral correctness is preserved.

10.2. Discussion on research questions

The above comparative analysis provides a basis to answer the research questions formulated in Section 1. With respect to RQ1, the previous discussion puts into evidence

[9] Configurative Process Modeling partially supports structural correctness only when customizing the model by restriction. Customization by extension in this approach does not guarantee correctness.

a number of core elements shared across all approaches. All approaches take as starting point a host process modeling language — usually a conceptual one rather than an executable one — on top of which a notion of *variation point* is added. Variation points are associated with specific model elements, which usually are control-flow elements (activities or gateways) but in some approaches can also be resources and objects.

On top of this common core, shared by all customizable process modeling approaches, we observe three key differentiating features. Firstly, some approaches allow variation points in a customizable model to be linked to elements in a domain model in order to assist the user during the customization of the model. Secondly, some approaches ensure that the customized models are structurally and behaviorally correct, disallowing combinations of customization options that would lead to an incorrect model. In three of the surveyed approaches, correctness is ensured at the expense of constraints on the structure of the models (block-structuredness), but in other cases, it is ensured for models with arbitrary topology. A third differentiating feature is given by the dichotomy between customization by restriction vs. by extension. While support for the former is widespread, the latter is only supported by a handful of approaches.

These distinguishing features constitute possible criteria for selecting an approach for a given purpose (cf. RQ2). If the set of variants of a given process is expected to grow incrementally after initial creation of a customizable process model, customization by extension is more convenient from a maintenance perspective. In this case, one starts with a customizable process model capturing the known variants. When a new variant is identified, its additional behavior with respect to the existing customizable model can be added as an extension point if it is well-confined. In approaches that only support customization by restriction, each new variant requires one to update the customizable process model, because the customizable model captures the union of all variants. In contrast, when configuration by extension is used, the customizable process model may capture only a core subset of the behavior of the variants. Variant-specific behavior can be captured in the extension points.

Meanwhile, if the decisions required for customization are complex and interdependent, approaches that link the customizable process model to a domain model and that provide customization guidance are preferable. If in addition the customizable process model is large and complex, approaches that support correctness checking during configuration may prove most useful. In this respect, it is not surprising that approaches based on customization by restriction tend to put emphasis on correctness-checking and guidance. Indeed, as the customizable model becomes larger, updates to it become more error-prone, and approaches based on customization by restriction lead to larger models since these models need to capture the union of all variants.

With reference to RQ3, we note that only a handful of approaches offer guidance and iterative feedback to the user in the selection of customization options. The few approaches that offer such guidance focus on ensuring that each selected customization option satisfies the domain constraints, or that the customized process model is correct. However, they do not address the question of which option (among those that are feasible) can lead to a customized process model with better performance with respect to relevant process performance measures. In other words, the relation between customization and business process performance has been so far neglected.

Second, whilst it is positive to observe that about half of the approaches have been implemented and at least partially validated in one way or another, there is a relative scarcity of comparative empirical evaluations. Barring two comparative studies [Torres et al. 2013; Döhring et al. 2014] focusing on a couple of approaches, there is a lack of evidence to back any statement that one customizable process modeling approach is more usable than others in a particular setting. This lack of comparative evaluations is arguably a major gap in the state of the art.

Third, we observe a lack of discussion on the question of how to construct a customizable process model in the first place, and how to maintain this artefact over time. It is generally assumed that a modeler will manually design the customizable process model using techniques similar to those employed to design classical (non-customizable) process models. Yet, given that a customizable process model represents an entire family of processes, the amount of information required to design such a model is usually an order of magnitude larger than that required to design a model of one singular process. This observation calls for the development of methods to assist process modelers during the design and update of customizable models.

Initial research on the design of customizable process models has led to algorithms for constructing a customizable process model from a collection of separate models of process variants [La Rosa et al. 2013; Assy et al. 2014], as well as algorithms for constructing a customizable process model from event logs extracted from enterprise systems [Buijs et al. 2013; Ekanayake et al. 2013]. Another approach is to extract a "common" (reference) process model out of a collection of models of process variants [Li et al. 2009]. This reference model may be particularly suited as a starting point for "customization by extension" approaches, since the reference model captures the "greatest common denominator" between existing variants, and the variant-specific behavior can then be added via extension points.

11. RELATED WORK

Ayora et. al. [Ayora et al. 2014] conducted a systematic literature review to evaluate existing variability support across all stages of the business process lifecycle. The authors considered 63 primary studies based on eight research questions (such as underlying business process modeling language used, tools available for enabling process variability, validation of methods proposed). Based on their findings, they developed the VIVACE framework to enable process engineers to evaluate existing process variability approaches. They then evaluate three approaches in depth using their framework: C-EPCs, Provop and PESOA. Our survey differs as we focus on the customization (configuration) of process models rather than approaches dealing with one or several phases of the lifecycle. Also, our comparison covers a superset of the above approaches.

Valenca et al. [Valenca et al. 2013] presented a literature mapping study in the field of business process variability covering 80 publications. Their objectives were to identify characteristics of business process variability (including design-time and run-time variability); to identify available approaches for business process variability management; and to identify challenges in this field. In line the nature of literature mapping studies, their study lists and categorizes a wide range of approaches, but without analyzing and comparing them in detail. In contrast, the present survey describes each main approach in detail, applies it to an example, and includes a comparative analysis. Further, our search returned a superset of the publications in [Valenca et al. 2013]. The latter remark also applies to a shorter and less systematic survey by Mechrez and Reinhartz-Berger [Mechrez and Reinhartz-Berger 2014].

A number of systematic reviews within the related domain of Software Product Line Engineering (SPLE) have been conducted. For instance, Chen et al. [Chen et al. 2009] conducted a systematic review of variability management in SPLE that included 33 papers. Their purpose was to provide an overview of different aspects of variability management approaches such as scalability and product derivation. Another one is by dos Santos Rocha and Fantinato [dos Santos Rocha and Fantinato 2013]. They conducted a systematic literature review to assess Software Product Line (SPL) approaches for BPM. Having reviewed 63 papers, they conclude that SPL approaches for BPM, while it is gaining maturity, are still at an inadequate level. Benavides et al. [Benavides et al. 2010] conducted a comprehensive literature review covering 53 papers to investigate existing proposals of automated analysis of feature models (within the context of

SPLE). Finally, Chen and Babar [Chen and Babar 2011] performed a systematic literature review that resulted in 97 papers being closely examined for assessing the status of evaluation of variability management approaches within SPLE. These reviews had a different objective and as such, cover other aspects of variability management as compared to our survey, namely understandability and maturity of evaluation. While they all contribute valuable insights, they share the commonality of being focused on the domain of SPLE. Furthermore, Chen et al. [Chen et al. 2009; Chen and Babar 2011] considered variability management within SPLE but dos Santos Rocha and Fantinato [dos Santos Rocha and Fantinato 2013] and Benavides et al. [Benavides et al. 2010] did not focus on variability management in particular. Our survey distinguishes itself from the above ones in that it focuses on variability management and second, it focuses on business processes as the artifact for which variability is captured and exploited. The distinctness of our survey with respect to the above ones is confirmed by the fact that the overlap of primary study papers is limited to a maximum of five papers – the overlap between dos Santos Rocha and Fantinato and ours is 5, Chen et al. and ours is 2 and for Benavides et al. it is 1.

Finally, [Torres et al. 2013] and [Döhring et al. 2014] compared a subset of the approaches reviewed in the present survey using different evaluation lenses. Torres et al. [Torres et al. 2013] compared C-EPCs and Provop in terms of understandability, based on a cognitive psychology framework. Döhring et al. [Döhring et al. 2014] conducted an empirical evaluation to assess the maintainability of process model variants in C-YAWL vs. vBPMN in terms of modularization support and customization type (i.e. restriction vs. restriction + extension). These papers examine non-functional aspects not considered in our survey, and as such they are complementary.

Mili et al. [Mili et al. 2010] survey, categorize and summarize different modeling languages used to describe business processes, covering in particular the languages figuring in the “Process modeling language” column of Table XII (e.g. BPMN, UML ADs). Their survey however does not touch upon the question of how to capture process variability in general and design-time variability in particular. As such, the scope of the present survey is disjoint and complementary to the one by Mili et al.

12. CONCLUSION

This survey has put into evidence a wide heterogeneity of features and levels of sophistication across existing approaches to customizable process models. Still, the survey has highlighted a common core shared by all of them and key differentiating features.

All approaches take as starting point a host process modeling language and add to it a notion of variation point. A variation point may be a modeling element that appears, does not appear or appears in one of multiple possible forms (customization by restriction) or a point in the process where additional behavior is allowed (configuration by extension). While virtually all approaches support customization by restriction, only a handful support customization by extension. Support for the latter constitutes one of the key differentiating features across the surveyed approaches and gives rise to a fundamental tradeoff that potential users need to consider when selecting an approach for a given scenario.

On the one hand, customization by extension is more suitable from a maintenance perspective. It allows one to start with a model capturing a core set of variants of the process. The behavior of additional variants can then be added via extension points, especially if the additional behavior of these variants can be confined to specific points in the customizable process model. The price to pay however is that the customizable process model itself only captures a subset of the behavior of the variants – additional behavior remains somehow hidden behind the extension points.

On the other hand, customization by restriction is more suitable when the set of variants is stable, since every new variant or every change to an existing variant re-

quires an update to the customizable process model. Additionally, this approach leads to larger models, since the customizable model has to capture the union of all variants. The latter hinders maintainability. Still, the models used in customization by restriction give a full picture of all variants and their dependencies. Also, because the behavior of all variants is captured in the customizable model, customization by restriction lends itself better to correctness checking, something that can only be achieved in customization by extension at the price of constraining the set of allowed extensions and the places in the customizable process model where these extensions can be inserted.

Despite the breadth of literature in the field of customizable process modeling, we have noted two areas that remain underdeveloped. First, there is a lack of effective methods and tool support to assist users in the creation, use (in particular customization) and maintenance of these models. Surprisingly, there has been little research on these questions. A handful of automated approaches to automatically construct a customizable process model from a collection of variants have been proposed [Li et al. 2009; La Rosa et al. 2013; Assy et al. 2014]. However, the user is then left with the task of fine-tuning these models, linking them with domain models and subsequently maintaining them. In a similar vein, little attention has been given to the question of how to guide users during the customization of customizable process models. The lack of methods and tools to support the full lifecycle of customizable process models (creation, use and maintenance) might explain the very limited adoption of customizable process modeling languages in practice.[10]

Secondly, while about half of reviewed approaches have been validated (typically via case studies), there is a lack of comparative empirical evaluations with end users that would provide evidence to back any statement that one customizable process modeling approach is more usable than others in a particular setting. There is a case for shifting the focus in this field from the design of modeling approaches to the evaluation of existing ones [Torres et al. 2013; Döhring et al. 2014].

Looking forward, the widespread adoption of multi-tenant enterprise systems has opened the possibility to use customizable process models to drive the configuration of such systems. At present, the configuration of multi-tenant systems is manual and resource-intensive due to the large number of configuration points offered by such systems. Initial visions for multi-tenant system configuration based on customizable process models have been put forward [Fehling et al. 2011; van der Aalst 2011]. However, the realization and validation of these visions remain avenues for future research.

ACKNOWLEDGMENTS

We thank Arthur ter Hofstede for his valuable comments on early versions of this manuscript. This research is partly funded by the Australian Research Council (grant DP150103356) and the Estonian Research Council (grant IUT20-55).

REFERENCES

ACHER, M., COLLET, P., LAHIRE, P., AND FRANCE, R. B. 2010a. Composing Feature Models. In *Proceedings of SLE 2009*, M. van den Brand, D. Gašević, and J. Gray, Eds. Vol. 5969. Springer, 62–81.

ACHER, M., COLLET, P., LAHIRE, P., AND FRANCE, R. B. 2010b. Managing variability in workflow managing variability in workflow with feature model composition operators. In *9th International Conference on Software Composition (SC), Malaga, Spain*. Springer, 17–33.

ASSY, N., GAALOUL, W., AND DEFUDE, B. 2014. Mining configurable process fragments for business process design. In *Proc. of DESRIST*. Lecture Notes in Computer Science Series, vol. 8463. Springer, 209–224.

AYORA, C., TORRES, V., WEBER, B., REICHERT, M., AND PALECHANO, V. 2014. VIVACE: A framework for the systematic evaluation of variability support in process-aware information systems. *Information and Software Technology*.

[10]To the best of our knowledge, only the aEPCs approach is currently used in practice by NN Investment Partners to manage their process model collection, which counts some 500 models.

BACHMANN, F. AND CLEMENTS, P. C. 2005. Variability in software product lines. Technical report CMU/SEI-2005-TR-012.

BAIER, T., PASCALAU, E., AND MENDLING, J. 2010. On the suitability of aggregated and configurable business process models. In *Enterprise, Business-Process and Information Systems Modeling - 11th International Workshop, BPMDS 2010, and 15th International Conference, EMMSAD 2010, held at CAiSE 2010, Hammamet, Tunisia, Proceedings*, T. A. Halpin, J. Krogstie, S. Nurcan, E. Proper, R. Schmidt, and R. Ukor, Eds. Lecture Notes in Business Information Processing Series, vol. 50. 108–119.

BECKER, J., ALGERMISSEN, L., DELFMANN, P., AND NIEHAVES, B. 2006. Configurable reference process models for public administrations. In *Encyclopedia of Digital Government*. 220–223.

BECKER, J., DELFMANN, P., DREILING, A., KNACKSTEDT, R., AND KUROPKA, D. 2004. Configurative Process Modeling – Outlining an Approach to increased Business Process Model Usability. In *Proceedings of the 14th Information Resources Management Association International Conference*, M. Khosrow-Pour, Ed. IRM Press.

BECKER, J., DELFMANN, P., AND KNACKSTEDT, R. 2007a. Adaptive Reference Modeling: Integrating Configurative and Generic Adaptation Techniques for Information Models. In *Proceedings of the Reference Modeling Conference (RM'06)*, J. Becker and P. Delfmann, Eds. Springer, 27–58.

BECKER, J., JANIESCH, C., KNACKSTEDT, R., AND RIEKE, T. 2007b. Facilitating change management with configurative reference modelling. *International Journal for Information Systems and Change Management 2,* 1, 81–99.

BECKER, J., KNACKSTEDT, R., PFEIFFER, D., AND JANIESCH, C. 2007c. Configurative method engineering - on the applicability of reference modeling mechanisms in method engineering. In *Proc. of the 13th Americas Conference on Information Systems (AMCIS 2007)*. 1–12.

BENAVIDES, D., SEGURA, S., AND RUIZ-CORTS, A. 2010. Automated analysis of feature models 20 years later: A literature review. *Information Systems 35,* 6, 616–636.

BOUCHER, Q., PERROUIN, G., DEPREZ, J.-C., AND HEYMANS, P. 2012. Towards configurable iso/iec 29110-compliant software development processes for very small entities. In *Proceedings of the 19th European Conference "Systems, Software and Services Process Improvement" (EuroSPI 2012)*, D. Winkler, R. V. O'Connor, and R. Messnarz, Eds. Communications in Computer and Information Science Series, vol. 301. Springer, 169–180.

BUIJS, J. C. A. M., VAN DONGEN, B. F., AND VAN DER AALST, W. M. P. 2013. Mining configurable process models from collections of event logs. In *BPM*. Lecture Notes in Computer Science Series, vol. 8094. Springer-Verlag, Berlin, 33–48.

CHEN, L. AND BABAR, M. 2011. A systematic review of evaluation of variability management approaches in software product lines. *Information and Software Technology 53,* 4, 344–362.

CHEN, L., BABAR, M., AND ALIO, N. 2009. Variability management in software product lines: a systematic review. *7328*, 190–205.

CZARNECKI, K. AND ANTKIEWICZ, M. 2005. Mapping Features to Models: A Template Approach Based on Superimposed Variants. In *Proceedings of the 4th International Conference on Generative Programming and Component Engineering*, R. Glück and M. R. Lowry, Eds. Springer, 422–437.

CZARNECKI, K., HELSEN, S., AND EISENECKER, U. 2005. Formalizing Cardinality-Based Feature Models and Their Specialization. *Software Process: Improvement and Practice 10,* 1, 7–29.

CZARNECKI, K. AND PIETROSZEK, K. 2006. Verifying feature-based model templates against well-formedness OCL constraints. In *Proc. of Generative Programming and Component Engineering*. ACM, 211–220.

DAVIS, R. AND BRABANDER, E. 2007. *ARIS Design Platform: Getting Started with BPM*. Springer.

DELFMANN, P., JANIESCH, C., KNACKSTEDT, R., RIEKE, T., AND SEIDEL, S. 2006. Towards Tool Support for Configurative Reference Modeling – Experiences from a Meta Modeling Teaching Case. In *Proceedings of the 2nd International Workshop on Meta-Modelling (WoMM'06)*, S. Brockmans, J. Jung, and Y. Sure, Eds. LNI Series, vol. 96. GI, 61–83.

DELFMANN, P., RIEKE, T., AND SEEL, C. 2007. Supporting enterprise systems introduction by controlling enabled configurative reference modelling. In *Proceedings of the Reference Modeling Conference (RM'06)*, J. Becker and P. Delfmann, Eds. Springer, 79–102.

DÖHRING, M., REIJERS, H., AND SMIRNOV, S. 2014. Configuration vs. adaptation for business process variant maintenance: an empirical study. *Inf. Syst. 39*, 108–133.

DOS SANTOS ROCHA, R. AND FANTINATO, M. 2013. The use of software product lines for business process management: A systematic literature review. *Information and Software Technology 55,* 8, 1355–1373.

DREILING, A., ROSEMANN, M., VAN DER AALST, W., HEUSER, L., AND SCHULZ, K. 2006. Model-Based Software Configuration: Patterns and Languages. *European Journal of Information Systems 15,* 6, 583–600.

DREILING, A., ROSEMANN, M., VAN DER AALST, W., SADIQ, W., AND KHAN, S. 2005. Model-Driven Process Configuration of Enterprise Systems. In *Wirtschaftsinformatik 2005: eEconomy, eGovernment, eSociety*, O. K. Ferstl, E. Sinz, S. Eckert, and T. Isselhorst, Eds. Physica–Verlag, 687–706.

EKANAYAKE, C., DUMAS, M., GARCÍA-BAÑUELOS, L., AND LA ROSA, M. 2013. Slice, Mine and Dice: Complexity-Aware Automated Discovery of Business Process Models. In *Business Process Management*. Lecture Notes in Computer Science Series, vol. 8094. Springer, 49–64.

FEHLING, C., LEYMANN, F., SCHUMM, D., KONRAD, R., MIETZNER, R., AND PAULY, M. 2011. Flexible process-based applications in hybrid clouds. In *Proceedings of the IEEE International Conference on Cloud Computing (CLOUD)*. IEEE, 81–88.

FETTKE, P. AND LOOS, P. 2003. Classification of Reference Models - A Methodology and its Application. *Information Systems and e-Business Management 1,* 1, 35–53.

GEORGAKOPOULOS, D., HORNICK, M., AND SHETH, A. 1995. An Overview of Workflow Management: From Process Modeling to Workflow Automation Infrastructure. *Distributed and Parallel Databases 3*, 119–153.

GOTTSCHALK, F., VAN DER AALST, W., AND JANSEN-VULLERS, M. 2007. SAP WebFlow Made Configurable: Unifying Workflow Templates into a Configurable Model. In *International Conference on Business Process Management (BPM 2007)*, G. Alonso, P. Dadam, and M. Rosemann, Eds. Lecture Notes in Computer Science Series, vol. 4714. Springer-Verlag, Berlin, 262–270.

GOTTSCHALK, F., VAN DER AALST, W., JANSEN-VULLERS, M., AND LA ROSA, M. 2008. Configurable Workflow Models. *International Journal of Cooperative Information Systems 17,* 2, 177–221.

GOTTSCHALK, F., WAGEMAKERS, T., JANSEN-VULLERS, M., VAN DER AALST, W., AND LA ROSA, M. 2009. Configurable Process Models: Experiences From a Municipality Case Study. In *Proc. of CAiSE*, P. van Eck, J. Gordijn, and R. Wieringa, Eds. Lecture Notes in Computer Science Series, vol. 5565. Springer-Verlag, Berlin, 486–500.

HALLERBACH, A., BAUER, T., AND REICHERT, M. 2008. Managing Process Variants in the Process Life Cycle. In *10th International Conf. on Enterprise Information Systems (ICEIS'08)*. Vol. 22. 154–161.

HALLERBACH, A., BAUER, T., AND REICHERT, M. 2009a. Guaranteeing Soundness of Configurable Process Variants in Provop. In *CEC*. IEEE, 98–105.

HALLERBACH, A., BAUER, T., AND REICHERT, M. 2009b. Issues in Modeling Process Variants with Provop. In *Business Process Management 2008 Workshops*, D. Ardagna, M. Mecella, and J. Yang, Eds. Lecture Notes in Business Information Processing Series, vol. 17. Springer.

HALLERBACH, A., BAUER, T., AND REICHERT, M. 2010. Capturing variability in business process models: The provop approach. *Journal of Software Maintenance and Evolution: Research and Practice 22,* 6-7, 519–546.

HEVNER, A., MARCH, S., PARK, J., AND RAM, S. 2004. Design Science in Information Systems Research. *MIS Quarterly 28,* 1, 75–105.

KUMAR, A. AND YAO, W. 2009. Process materialization using templates and rules to design flexible process models. In *RuleML*, G. Governatori, J. Hall, and A. Paschke, Eds. Lecture Notes in Computer Science Series, vol. 5858. Springer, 122–136.

KUMAR, A. AND YAO, W. 2012. Design and management of flexible process variants using templates and rules. *Computers in Industry 63,* 2, 112–130.

LA ROSA, M., DUMAS, M., TER HOFSTEDE, A., AND MENDLING, J. 2011. Configurable Multi-Perspective Business Process Models. *Information Systems 36,* 2, 313–340.

LA ROSA, M., DUMAS, M., UBA, R., AND DIJKMAN, R. M. 2013. Business process model merging: An approach to business process consolidation. *ACM Transactions on Software Engineering Methodology 22,* 2, 11.

LI, C., REICHERT, M., AND WOMBACHER, A. 2009. Discovering Reference Models by Mining Process Variants Using a Heuristic Approach. In *Business Process Management (BPM 2009)*, U. Dayal, J. Eder, J. Koehler, and H. Reijers, Eds. Lecture Notes in Computer Science Series, vol. 5701. Springer-Verlag, Berlin, 344–362.

LÖNN, C.-M., UPPSTRÖM, E., WOHED, P., AND JUELL-SKIELSE, G. 2012. Configurable process models for the swedish public sector. In *Proceedings of the 24h International Conference on Advanced Information Systems Engineering (CAiSE 2012)*, J. Ralyté, X. Franch, S. Brinkkemper, and S. Wrycza, Eds. Lecture Notes in Computer Science Series, vol. 7328. Springer, 190–205.

MECHREZ, I. AND REINHARTZ-BERGER, I. 2014. Modeling design-time variability in business processes: Existing support and deficiencies. In *Proc. of BPMDS 2014 and EMMSAD 2014*. Springer, 378–392.

MILI, H., TREMBLAY, G., JAOUDE, G. B., LEFEBVRE, E., ELABED, L., AND EL-BOUSSAIDI, G. 2010. Business process modeling languages: Sorting through the alphabet soup. *ACM Computing Surveys 43,* 1, Article 4.

MOON, M., HONG, M., AND YEOM, K. 2008. Two-level variability analysis for business process with reusability and extensibility. In *Proceedings of the 32nd Annual IEEE International Computer Software and Applications Conference (COMPSAC), Turku, Finland*. IEEE Computer Society, 263–270.

PESIC, M., SCHONENBERG, H., AND VAN DER AALST, W. 2007. DECLARE: Full Support for Loosely-Structured Processes. In *Proceedings of the Eleventh IEEE International Enterprise Distributed Object Computing Conference (EDOC 2007)*, M. Spies and M. Blake, Eds. IEEE Computer Society, 287–298.

PUHLMANN, F., SCHNIEDERS, A., WEILAND, J., AND WESKE, M. 2005. Variability Mechanisms for Process Models. PESOA-Report TR 17/2005, Process Family Engineering in Service-Oriented Applications (PESOA). BMBF-Project. 30 June.

REICHERT, M. AND DADAM, P. 1998. ADEPT$_{flex}$: Supporting Dynamic Changes of Workflow without Loosing Control. *Journal of Intelligent Information Systems 10,* 2, 93–129.

REICHERT, M. AND WEBER, B. 2012. *Enabling Flexibility in Process-Aware Information Systems: Challenges, Methods, Technologies*. Springer-Verlag, Berlin.

REIJERS, H., MANS, R., AND VAN DER TOORN, R. 2009. Improved Model Management with Aggregated Business Process Models. *Data Knowl. Eng. 68,* 2, 221–243.

REINHARTZ-BERGER, I., SOFFER, P., AND STURM, A. 2009. Organizational Reference Models: Supporting an Adequate Design of Local Business Processes. *International Journal of Business Process Integration and Management 4,* 2, 134–149.

REINHARTZ-BERGER, I., SOFFER, P., AND STURM, A. 2010. Extending the Adaptability of Reference Models. *IEEE Transactions on Systems, Man and Cybernetics part A 40,* 5, 1045–1056.

REINHARTZ-BERGER, I. AND STURM, A. 2007. Enhancing UML Models: A Domain Analysis Approach. *Journal on Database Management (special issue on UML Topics) 19,* 1, 74–94.

RINDERLE, S., REICHERT, M., AND DADAM, P. 2004. Correctness Criteria For Dynamic Changes in Workflow Systems: A Survey. *Data and Knowledge Engineering 50,* 1, 9–34.

ROSEMANN, M. 2003. Application Reference Models and Building Blocks for Management and Control (ERP Systems). In *Handbook on Enterprise Architecture*, P. Bernus, L. Nemes, and G. Schmidt, Eds. Springer, 596–616.

ROSEMANN, M. AND VAN DER AALST, W. 2003. A Configurable Reference Modelling Language. BPM Center Report BPM-03-08, BPMcenter.org. (Later published as ROSEMANN, M. AND AALST, W. 2007).

SADIQ, S., ORLOWSKA, M., AND SADIQ, W. 2005. Specification and validation of process constraints for flexible workflows. *Inf. Syst. 30,* 5, 349–378.

SADIQ, S., SADIQ, W., AND ORLOWSKA, M. 2001. Pockets of Flexibility in Workflow Specification. In *Proceedings of the 20th International Conference on Conceptual Modeling (ER 2001)*. Lecture Notes in Computer Science Series, vol. 2224. Springer-Verlag, Berlin, 513–526.

SCHNIEDERS, A. 2006. Variability Mechanism Centric Process Family Architectures. In *Proceedings of the 13th IEEE International Symposium and Workshop on Engineering of Computer Based Systems (ECBS)*, M. Riebisch, P. Tabeling, and W. Zorn, Eds. IEEE Computer Society, 289–298.

SCHNIEDERS, A. AND PUHLMANN, F. 2006. Variability Mechanisms in E-Business Process Families. In *Proceedings of the 9th International Conference on Business Information Systems (BIS'06)*, W. Abramowicz and H. Mayr, Eds. LNI Series, vol. 85. GI, 583–601.

SVAHNBERG, M., VAN GURP, J., AND BOSCH, J. 2005. A taxonomy of variability realization techniques. *Softw., Pract. Exper. 35,* 8, 705–754.

TORRES, V., ZUGAL, S., WEBER, B., REICHERT, M., AYORA, C., AND PELECHANO, V. 2013. A qualitative comparison of approaches supporting business process variability. In *Business Process Management Workshops*. Lecture Notes in Business Information Processing Series, vol. 132. Springer, 560–572.

VALENCA, G., ALVES, C., ALVES, V., AND NIU, N. 2013. A systematic mapping study on business process variability. *Int. Journal of Computer Science & Information Technology 5,* 1, 1–21.

VAN DER AALST, W., DREILING, A., GOTTSCHALK, F., ROSEMANN, M., AND JANSEN-VULLERS, M. 2006. Configurable Process Models as a Basis for Reference Modeling. In *Proceedings of the Business Process Management 2005 Workshops*, E. Kindler and M. Nüttgens, Eds. Springer, 76–82.

VAN DER AALST, W., HEE, K., TER HOFSTEDE, A., SIDOROVA, N., VERBEEK, H., VOORHOEVE, M., AND WYNN, M. 2011. Soundness of Workflow Nets: Classification, Decidability, and Analysis. *Formal Aspects of Computing 23,* 3, 333–363.

VAN DER AALST, W., LOHMANN, N., AND LA ROSA, M. 2012. Correctness Ensuring Process Configuration: An Approach Based on Partner Synthesis. *Information Systems 37,* 6, 574–592.

VAN DER AALST, W. M. P. 2011. Business process configuration in the cloud: How to support and analyze multi-tenant processes? In *Proceedings of the 9th IEEE European Conference on Web Services (ECOWS)*. IEEE, 3–10.

WEBER, B., REICHERT, M., AND RINDERLE-MA, S. 2008. Change Patterns and Change Support Features: Enhancing Flexibility in Process-Aware Information Systems. *Data and Knowledge Engineering 66,* 3, 438–466.

ELECTRONIC APPENDIX

The electronic appendix for this article can be accessed in the ACM Digital Library.

Online Appendix to: Business Process Variability Modeling: A Survey

MARCELLO LA ROSA, Queensland University of Technology, Australia
WIL M.P. VAN DER AALST, Eindhoven University of Technology, The Netherlands
MARLON DUMAS, University of Tartu, Estonia and Queensland University of Technology, Australia
FREDRIK P. MILANI, University of Tartu, Estonia

A. SEARCH PROTOCOL

The literature search followed the principles of systematic literature review in [Kitchenham 2004]. As proposed in [Webster and Watson 2002], the first step is to define the aim of the search, which is to identify a relatively complete list of studies that propose customizable process models to manage business process variability. To ensure that every important study was found, we applied several search strategies, as recommended in [Fink 2010; Okoli and Schabram 2010; Randolph 2009; Levy and Ellis 2006; Kitchenham 2004]. The primary search was done with a well-known electronic literature database as proposed by several studies [Fink 2010; Okoli and Schabram 2010; Randolph 2009; Rowley and Slack 2004]. We used Google Scholar, which encompasses all relevant databases such as ACM Digital Library and IEEE Xplore. We also extended the search by using complementary key terms, which we found in the papers returned by the primary search. Finally existing mappings related to business process variability were also examined to ensure that all relevant studies had been identified in our search.

A.1. Search string development

In the primary search, we used key terms related to business process variability and customizable process models. We first determined that the term "customization" is associated with "variation" and "configuration". Accordingly, we constructed queries by combining the keyword "business process" with "customization", "customizability", "customizable", "variation", "variability", "configuration", "configurability", and "configurable". Each search was done separately using the conjunction of "business process" with each of the above terms, resulting in eight search strings. We also noted that the keyword "flexibility" is often associated with customization, and thus also constructed queries combining "business process" with "flexible" and with "flexibility", leading to two further search strings. Since the term "workflow" is sometimes used as a quasi-synonym of "business process", we also included queries combining "workflow" with the above keywords in the same manner (resulting in ten further strings). As we conducted the search, we noted additional terms appearing in the titles of relevant papers, namely "business process variant", "configurable reference model", "reference model adaptability", "reference model adaptation", "reference model flexibility" and "configurable EPC". We therefore conducted searches using these key terms, leading to six further search strings.

We are aware that extensive research pertaining to variability modeling in software systems - most notably using feature models [Schobbens et al. 2006] - has been conducted in the field of Software Product Line Engineering (SPLE) [Czarnecki and Eisenecker 2000]. Some of this research addresses the question of variability in business process models captured by means of UML activity diagrams. Accordingly, we also

DOI 10.1145/0000000.0000000 http://doi.acm.org/10.1145/0000000.0000000

included queries composed of "UML activity diagram" in conjunction with "software product line" or "feature model" (two further search strings).

For each of the 28 queries (one per search string), we gathered the first 100 hits in Google Scholar. Based on the title of the study, we filtered out papers that were clearly out of scope. We observe that running these 28 queries in isolation, rather than running a single query with the disjunction of all the identified search strings, led to a much larger pool of papers to analyze. The queries were run in August 2015. However, in order to have a consistent snapshot, we restricted the search to return studies published until 2014. This resulted in over 2,400 hits.

A.2. Validation

In order to validate the choice of Google Scholar as the search database, we searched the following popular academic databases using the same search strings: ACM Digital Library, IEEE Xplore, ScienceDirect, Citeseer, Inspec, EI Compendex, SCOPUS and SpringerLink, as identified in [Kitchenham 2004]. We noted that these searches did not return any paper that was not already discovered by our primary search. Thus, we concentrated on the results returned by Google Scholar.

Kitchenham [Kitchenham 2004] recommends to validate trial search strings against lists of already known primary studies. Accordingly, we examined two existing literature mapping studies in the field of business process variability, namely [Ayora et al. 2014] and [Valenca et al. 2013]. [Ayora et al. 2014] develops an evaluation framework for variability management approaches, along the whole BPM lifecycle. We extracted from this mapping study publications related to design-time variability management via customizable process models. Out of the 63 primary studies identified in [Ayora et al. 2014], we found that all 23 studies which fall under the scope of our survey, were also retrieved by our search. Similarly, the mapping study of [Valenca et al. 2013] provides an inventory of 80 publications covering both design-time and run-time variability. We verified that all relevant publications were also found by our search. We also noted that our search returned several publications not covered in [Ayora et al. 2014] and [Valenca et al. 2013].

A.3. Study selection

After filtering out papers that were clearly out of scope (based on title), we proceeded to removing duplicates. As suggested by [Fink 2010; Okoli and Schabram 2010; Randolph 2009; Torraco 2005], we defined inclusion and exclusion criteria in order to ensure an unbiased selection of relevant studies. The development of criteria for inclusion and exclusion, as recommended in [Kitchenham 2004], was based on the objective and the scope of this survey. The assessment of each study against the inclusion and exclusion criteria was performed independently by two authors of this paper. The results were compared in order to resolve inconsistencies with the mediation of a third author.

The inclusion criteria for the first screening of results were:

(1) Does the study propose a method to either model or maintain a family of process model variants via a customizable process model?
(2) Does the study propose an approach to customize a customizable process model?
(3) Does the study have at least 10 citations?
(4) Is the paper at least 5 pages, single column or 3 pages, double column?

Each study for which the answer to all the above questions (inclusion criteria) was positive, was included. As such papers with less than the citation and length thresholds were excluded. The page limit was set because a short paper would not contain enough information for an evaluation. This initial filtering reduced the number of candidate papers to 370.

For example, the approaches presented in [Heuer et al. 2013; Asadi et al. 2014] were not included in this survey as they did not reach a sufficient number of citations. Other approaches, such as [Rastrepkina 2010; Meerkamm 2010], were excluded because they do not use customizable process models to manage a family of process model variants (e.g. [Meerkamm 2010] proposes to organize process model variants in a tree).

We then continued the screening process with an inspection of the abstract of each paper in order to exclude papers that fell within related but clearly distinct areas of study. In this part we used exclusion criteria. If the answer to any of the questions defining the exclusion criteria was positive, the paper was excluded from the survey. The exclusion criteria were:

(1) Does the study concern managing process model variants only at run-time (or process flexibility)?
(2) Does the study concern managing process model variants for exception handling?
(3) Does the study concern managing process model variants as automated workflow composition?

For example, the approach in [Lu et al. 2009] was excluded as it proposes to use domain constraints to adapt a process model instance (defined through a template) in order to derive a variant of such instance at run-time, which is then executed in a supporting system.

At the end of the search process, we obtained 66 relevant publications. In general, a given approach is covered by multiple publications. We also found that some approaches are subsumed by other approaches, i.e., the features and concepts in one approach are contained by another. In total, we found that the 66 publications cover 23 different approaches, out of which eleven main approaches subsume the other twelve approaches. Accordingly, this survey distinguishes eleven main approaches and twelve subsumed approaches.

The publications covered by the survey are listed in a supplemental spreadsheet available at `https://goo.gl/OAGzdj`. For each approach, the table identifies a primary (earliest) publication describing the approach and where available, additional publications describing further aspects of the same approach.

B. SUBSUMED APPROACHES

Twelve proposals for customizable process modeling are subsumed by the eleven main approaches described in this survey paper. We say that an approach A is subsumed by B if A supports a subset of the variability concepts of B. The focus is on the supported variability concepts and not on the process modeling language, supporting techniques or extra-functional aspects. Thus, for example, an approach A can be subsumed by B even if A and B are applied to different process modeling languages (e.g. EPC vs. UML ADs). Subsumed approaches are minor approaches compared to the approaches presented in this paper. They are briefly reviewed below.

B.0.1. Subsumed by C-iEPCs. The initial incarnation of the KobrA (Component-based Application Development) method [Atkinson et al. 2000] provides a mechanism to capture a family of process variants via a customizable process model. The purpose is that of customizing component-based software systems. As such, process models are employed for the description of components' behavior. Customization is done using UML ADs and is driven by a decision table (see Appendix C.3). Similar to C-iEPCs, XOR gateways in UML ADs can be marked as configurable (using a black background and a letter "M") to indicate that subsequent activities are optional. However, a transformation algorithm is not discussed. Further, there is no method to guarantee the correctness of the customized model, nor is there a mechanism to exclude unfeasible customized models as a result of wrong combinations of customization options. The

Kobra method has been implemented in a tool and validated in numerous industrial settings [Atkinson et al. 2008], though there is no information on the involvement of domain experts.

Korherr & List [Korherr and List 2007] present an approach which extends UML ADs with stereotypes to capture variability. In their approach, variability can be defined at the level of an atomic activity, a group of activities or an XOR gateway. An activity or group thereof can be defined as being `<<mandatory>>` (the activity or group must be retained during customization) or `<<optional>>` (the activity or group can be excluded during customization). An XOR gateway can be defined as being an `<<alternative choice>>` with a `0..1` range (at most one outgoing sequence flow must be selected during customization) or with a `1..*` range (at least one outgoing flow must be selected). It is also possible to state that the selection of an element (activity, group or flow) during customization requires the selection of another element elsewhere (denoted by a dependency arrow marked with the `<<requires>>` stereotype), or that the selection of an element excludes the selection of another one elsewhere in the model (denoted by an arrow with the `<<excludes>>` stereotype). Further constraints between configurable nodes can be defined using the Object Constraint Language (OCL). Abstraction from the customization of the process model can be achieved via the use of a UML Profile for variability, provided by this approach. This is similar to a feature model in terms of functionality (cf. Section C.1).

B.0.2. Subsumed by Configurable Workflows. A CoSeNet (Configurable Service Net) [Schunselaar et al. 2011; 2012] is an alternative representation of a configurable workflow model via a directed acyclic graph. CoSeNets have been designed to fulfill two requirements: i) always yield correct customized models and ii) being reversible, i.e. the initial process variants used to create the CoSeNet should be obtained through customization. Each leaf of a CoSeNet represents a process activity and each parent node represents a control-flow operator. The available operators are sequence, the gateways OR, AND, data-driven and event-driven XOR, and the structured REPEAT-UNTIL loop. Connections between nodes are achieved via special nodes, called VOID nodes, which are linked to parent and child nodes via arcs and do not bear any semantics. For example, an OR between "Prepare film for editing" and "Prepare tape for editing" means that either or both of these activities can be executed. A CoSeNet thus captures a block-structured process model where each single-entry single-exist fragment is identified by an operator and its children nodes. This structure guarantees the behavioral correctness of the process model by construction, since split and join within a fragment are of the same type. Customization is achieved by applying the hiding and blocking operations of Configurable Workflows to the VOID nodes. CoSeNets have been defined formally using the YAWL semantics, though a definition of the transformation algorithm is not available. A mapping from CoSeNets to plain YAWL models can be used for executing the configured models. However, the approach abstracts from data and resource aspects, thus effectively offering limited support for execution. Moreover, a mapping in the opposite direction is not described. This approach has been implemented via various plugins for the ProM environment[11] and used to capture process models from various municipalities.

B.0.3. Subsumed by aEPCs. Gröner et al. [Gröner et al. 2013] propose an approach similar to aEPCs. They rely on two artifacts: a plain block-structured BPMN model, which captures all variants of a business process family, and a feature model (cf. Section C.1) which captures the variability of the process domain. The two artifacts are linked by mapping features into BPMN activities, similar to the mapping of product hierarchies into EPC elements in the aEPCs approach. Customization is driven by fea-

[11] See `http://processmining.org`.

ture selection, such that those process model activities whose features are not selected, are removed from the process model. Hence, process models are customized by restriction of process behavior, and abstraction from the process model level is achieved via the use of feature models. No other BPMN element is customizable besides activities. The focus of this approach, however, is not on customization per se, but rather on ensuring consistency between the domain constraints defined on features (e.g. an XOR between two features), and the structural constraints of the process model imposed by its control-flow relations (e.g. an XOR-split between two activities). This is achieved by mapping both feature constraints and control-flow constraints into Description Logic and reasoning at the level of a single set of Description Logic constraints—an idea also explored in [La Rosa 2009] in the context of C-iEPC models. Guidance is only partially fulfilled, as the approach by Grner et al. prevents users from taking inconsistent customization decisions, but does not guide them in making the right decisions (e.g. via recommendations). All aspects of the approach by Grner et.al. are formalized, but given the focus on consistency checking, a transformation algorithm is not defined. For the same reason, correctness preservation is not discussed, leaving room for interpretation. In fact, even if the starting BPMN model is block-structured and syntactically correct, removal of activities may cause syntactical errors such as disconnections in the resulting BPMN model. The approach has been implemented in a tool (not publicly available) and validated using an e-store and a post-production scenario, though without involving domain experts.

B.0.4. Subsumed by PESOA. Razavian & Khosravi [Razavian and Khosravi 2008] propose an approach to define customizable process models in the form of UML ADs extended with a fixed set of stereotypes. The set of stereotypes is a subset of that in PESOA. There are two types of variation points: *optional* and *alternative*. An optional variation point allows the selection of at most one variant among the available ones; an alternative one allows the selection of exactly one variant. Variation points can be defined on both control-flow elements and on data objects. Specifically, an XOR-split can be marked with `<<opt_vp>>` to indicate an optional variation point, in which case the gateway is customized by choosing at most one of its outgoing flows, and with `<<alt_vp>>` to indicate an alternative variation point, in which case the gateway is customized by choosing exactly one of its outgoing flows. However, other types of gateways such as AND gateways cannot be customized. Activities can be marked as `<<optional>>` or as `<<vp_al>>`. In the former case, the activity can be excluded during customization, while in the latter case it can be customized to one of its variants. Interdependencies between model elements cannot be defined beyond stating that a variation point is either optional or alternative. As a result, only simple configuration scenarios can be captured. The customization of input and output data objects and data stores (used to persist data beyond a process instance) is achieved in the same way as for activities. It is also possible to mark a sub-process activity with the stereotype `<<variable>>` to indicate that the underlying model includes some variation points. The authors recognize that if no variant is chosen for an optional variation point in the control flow, the model may become disconnected. However the correctness of the customized model is not guaranteed.

A similar approach is provided by Ciuskys & Caplinskas [Ciuksys and Caplinskas 2007]. In this approach, only the activities of a UML AD can be defined as variation points (called *generic* activities). During customization, a generic activity can be replaced by one of several possible concrete (*non-generic*) activities. Alternatively, an activity may be removed during customization if it is marked as *optional*. The space of customization options is specified using a feature model (cf. Section C.1), where each feature corresponds to a (generic or non-generic) activity and where feature interdependencies can be defined. The features that are inner nodes in the feature model

represent generic activities, while the leaf features correspond to non-generic activities. One may customize a process model by selecting features in the feature model. These features then determine how the generic activities in the process model are customized. A Description Logic reasoner is used for checking the consistency of a given customization, expressed as a subset of features selected from the feature model.

Kulkarni & Barat [Kulkarni and Barat 2011] put forward a similar approach in the context of BPMN models. Here a generic activity (called *abstract activity*) can be replaced by a single (atomic) *concrete activity* or by an entire sub-process (called *composite activity*). Also abstract events can be replaced by concrete events. Gateways cannot be customized. Kulkarni & Barat suggest that feature models (cf. Section C.1) could be used to guide the customization process, but they do not specify any concrete mechanism for linking a process model with a feature model. Thus, effectively they do not provide any decision support for process model customization. A formalization of the basic notions is provided, though a transformation algorithm is not defined. The approach has not been implemented nor validated.

B.0.5. Subsumed by BPFM. Ripon et al. [Ripon et al. 2010] present an approach similar to BPFM using UML ADs. An activity marked with a stereotype `<<variant>>` represents a variation point and is linked to an entry in a *variant model* listing all possible options (i.e. variants) for customizing the activity. Such variants are also summarized in a decision table (see Appendix C.3) that is presented to the user. Multiple variants can be selected for the same variation point, depending on the constraints specified among the variants of the same variation point, though it is not clear how multiple variants, when selected, will be represented in the customized model. By selecting/deselecting variants from the decision table, one can determine which variant(s) of an activity will be picked during customization. While multiple variants can be selected for a variation point, different than BPFM, a cardinality cannot be specified. Further, the approach only works by restriction of variants.

Nguyen et al. [Nguyen et al. 2011] operate in the context of BPMN models. In this approach, variation points can be defined in BPMN activities, objects, as well as message flows connecting activities in different pools. Each variation point is assigned one or more variants and a minimum and maximum cardinality is attached to define the number of variants that can be selected. Dependencies between variants, within and across variation points can also be defined. The approach works on BPMN models at the conceptual level. Abstraction is achieved via the use of feature models (see Section C.1). An implementation of the approach as an Eclipse plugin is available.

B.0.6. Subsumed by Provop. vBPMN (variant BPMN) [Döhring et al. 2011; Döhring et al. 2014] is a Provop-based approach for design- and run-time customizations of executable process models using BPMN. In our survey we focus on the aspects of this approach related to design-time customization. Accordingly, the approach applies structural adaptations to a base model defined in BPMN and annotated with adjustment points. These adjustment points are indicated with black diamonds on top of activities (called *adaptive activities*) to identify customizable activities, as well as intermediate events marked with a square bracket to delimit a fragment of the model (called *adaptive segment*) that can be customized. Patterns from an extensive catalogue, defined in the form of syntactically correct block structures, can be assigned to adaptive activities or inserted into an adaptive segment, to customize the model. Thus, the only possible operation is INSERT, as opposed to Provop which also allows DELETE, MOVE and MODIFY. As such, vBPMN is the only approach surveyed that does not provide customization by restriction. An advantage of using INSERT only is that if the base model and the fragments to insert are correct, it is not possible to generate an incorrect customized model by design. Adaptation rules for applying patterns to adaptive parts of the model are only specified for run-time settings (i.e. they are triggered af-

ter the execution of a particular event). As such, the approach does not provide any decision support for customizing the process model at design-time. vBPMN has been implemented in a tool based on the jBoss Drools execution engine (not publicly available), which allows the customized models to be executed. However, inconsistencies in the data dependencies caused by customization of the control flow are not detected and avoided. Thus this approach only offers partial support for execution. vBPMN has been validated on complex business processes from the municipality domain, though without involving experts from this domain [Döhring et al. 2014]. Moreover, an empirical evaluation of this approach in comparison with C-YAWL has been carried out by the authors [Döhring et al. 2014] (cf. Section 11).

Santos et al. [Santos et al. 2010] propose to customize BPMN models using non-functional requirements. Their approach is also similar to Provop, though they do not directly annotate the base model. Rather, they equip the base model with a list of variation points each indicating a fragment (delimited by a reference to two adjustment points in the model) or an individual model element (activity or resource) from the base model that is customizable. Each variation point is then assigned a list of variants, i.e. model fragments that can be inserted into the variation points. Other operations are the deletion of the fragment or element identified by the variation point, and the substitution of a fragment with another. For example, one can replace a sequence flow with an entire fragment, remove an activity or add a lane (the BPMN element for representing resources). Simple exclusion dependencies can be specified between variants. The authors propose to achieve abstraction by driving the customization of the base model through a list of *non-functional requirements*, i.e. domain aspects, that can be linked to the variants.

Machado et al. [Machado et al. 2011] propose to extend BPMN with two aspect-oriented constructs: i) the *pointcut* (similar to Provop's adjustment point), to be applied to a base model, and ii) the *advice*, a pre-defined process model fragment representing a variant that can be inserted, removed or replaced before, after or around a pointcut in the base model. These constructs are then linked to a feature model (cf. Section C.1) capturing product line variability. The customization of the base model is then carried out by configuring the feature model, abstracting from the process model itself. The customized BPMN models can then be instantiated using Haskell as the host language. However, potential inconsistencies in data dependencies caused by customization are not fixed. The approach is partly implemented in a tool, not publicly available.

C. TECHNIQUES FOR DECISION SUPPORT

In this section we report on three techniques that can be used to provide decision support during process model customization. Two such techniques, namely Feature Models and Decision Tables, offer *abstraction* from the customizable process model and its variation points when performing a customization. This is achieved by capturing variability at the level of the domain in which the process model has been constructed, in order to allow users to reason in terms of domain properties rather process modeling elements. This is especially useful when the customizable process model features many interdependent variation points, as one would expect in a realistic scenario. A third technique, namely Questionnaire Models, also offers *guidance* for the customization of process models. Guidance entails the provision of mechanisms to guide users in making the right customization decisions, for example in the form of recommendations, avoiding inconsistent or irrelevant decisions.

C.1. Feature Models

Feature models are a family of techniques originally conceived to describe variability in software product lines in terms of their features, and later applied to different domains.

Various feature modeling languages have been proposed [Schobbens et al. 2006] since feature models were first introduced as part of the FODA (Feature Oriented Domain Analysis) method [Kang et al. 1990].

A feature model is represented graphically by one or more feature diagrams. A feature diagram is a rooted tree with high-level features decomposed into sub-features. A *feature* captures a property of the domain under analysis, that is relevant to a stakeholder. Figure 15 shows a possible feature diagram for the picture post-production domain, using the notation proposed in [Batory 2005]. There are features related to the options available for shooting, type of editing, transfer and finish.

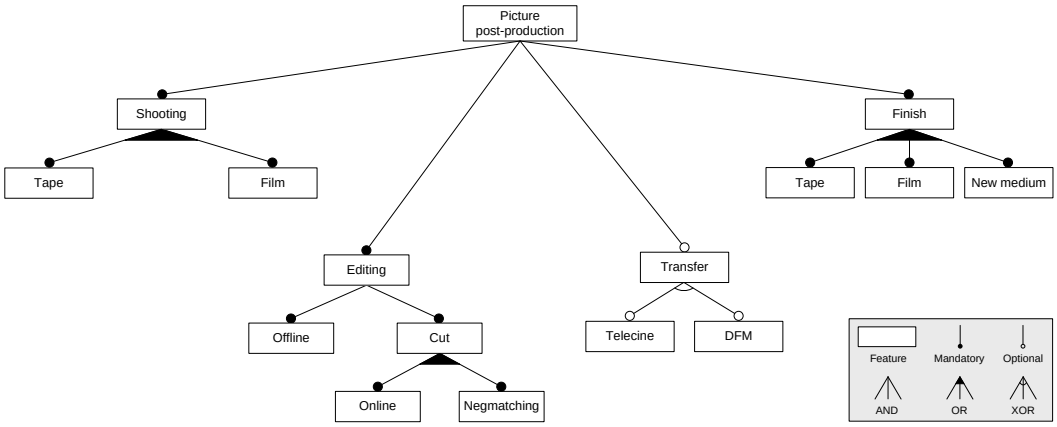

Fig. 15: A feature diagram for post-production.

A feature can be mandatory, or optional (in which case it can be deselected), and can be bound to other features via constraints. *Feature constraints* can be expressed as arbitrary propositional logic expressions over the values of features [Batory 2005]. For example, the sub-feature "Negmatching" of "Cut" must be deselected if the sub-feature "Film" of "Shooting" is not selected.

Feature constraints between the sub-features of a same feature can also be represented graphically. This way restrictions with respect to the number of sub-features a feature can have can be modeled. These relations can be: AND (all the sub-features must be selected), XOR (only one sub-feature can be selected) and OR (one or more can be selected). OR relationships can be further specified with an $n : m$ cardinality [Czarnecki et al. 2005], where n indicates the minimum and m indicates the maximum number of allowed sub-features. For example, in Figure 15 the sub-features of "Cut" are bound by an OR relation as it is possible to have more than one type of cut in post-production.

We observe that while an optional feature always represents an element of variability, a mandatory feature does not necessarily represent a commonality in the domain under analysis. In fact, a mandatory feature can still be excluded if it has an XOR/OR relation with its sibling features. This is the case of the sub-features of "Finish", which are all mandatory (a choice on the finish is required), though it is possible to choose any subset of them due to the OR relation.

A *feature configuration* specifies a valid scenario in terms of features selected/deselected, i.e. a scenario that complies with the feature constraints. Figure 16 depicts the feature diagram for post-production configured for a project shot on tape, edited online and delivered on film.

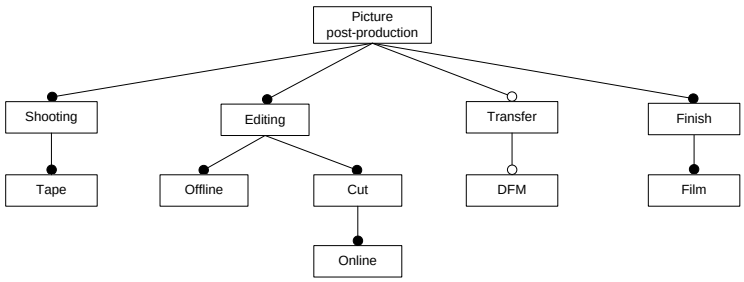

Fig. 16: A possible configuration for the post-production feature diagram.

Various tools supporting the definition and configuration of feature models are available. Some examples are the *AHEAD Tool Suite*,[12] the Eclipse plugin *FeatureIDE*,[13] and the online toolset *SPLOT*.[14]

Although the initial aim of feature models was to facilitate the configuration of software product families, this technique has also been used to provide abstraction for the customization of process models in various approaches. Feature models are used by PESOA (see Section 8.1), Superimposed Variants (cf. Section 7.2), Feature Model Composition (cf. Section 8.3), and the approaches by Ciuskys & Caplinskas (cf. Section B.0.4), by Nguyen et al. (cf. Section B.0.5), by Machado et al. (cf. Section B.0.6), and by Gröner et al. (cf. Section B.0.3). Korherr & List (cf. Section B.0.1) use UML Profiles for variability which are similar to feature diagrams.

In these approaches, one can customize a process model by selecting/deselecting features from a feature model. In order to do so, one has to first establish a mapping between features on the one hand, and variants of variation points in the process model on the other hand. Once a feature configuration has been completed, this mapping is used to automatically select the right variant(s) for each variation point of the customizable process model. Then a transformation algorithm, if available, is triggered to obtain the customized process model.

C.2. Questionnaire Models

Questionnaire models [La Rosa et al. 2009] are another technique for representing the variability of a domain. The idea is to organize a set of features, called *domain facts*, into *questions* that can be posed to users in order to configure the domain in question.

Figure 17 shows a possible questionnaire model for post-production, where all questions and facts have been assigned a unique identifier. Questions group domain facts according to their content, so that all the domain facts of a same question can be set at once by answering the question. For example, the question “What type of shooting has been used?” groups the domain facts “Tape shooting” and “Film shooting”. Each domain fact is a boolean variable and has a default value, which can be used to identify the most common choice for that fact. For example, since the majority of production projects are shot on tape because it is less expensive than film, we can assign a default value of true to “Tape shooting”, and of false to “Film shooting”. Moreover, a domain fact can be marked as mandatory if it needs to be explicitly set when answering the questionnaire. If a non-mandatory fact is left unset, i.e. if the corresponding question is left unanswered, its default value can be used to answer that question. In this way, each domain fact will always be set, either explicitly by an answer or by using

[12] See `http://www.cs.utexas.edu/users/schwartz/ATS.html`.
[13] See `http://fosd.de/fide`.
[14] See `http://www.splot-research.org`.

its default value. Accordingly, the mandatory attribute of a domain fact has different semantics than the mandatory attribute of a feature in a feature model.

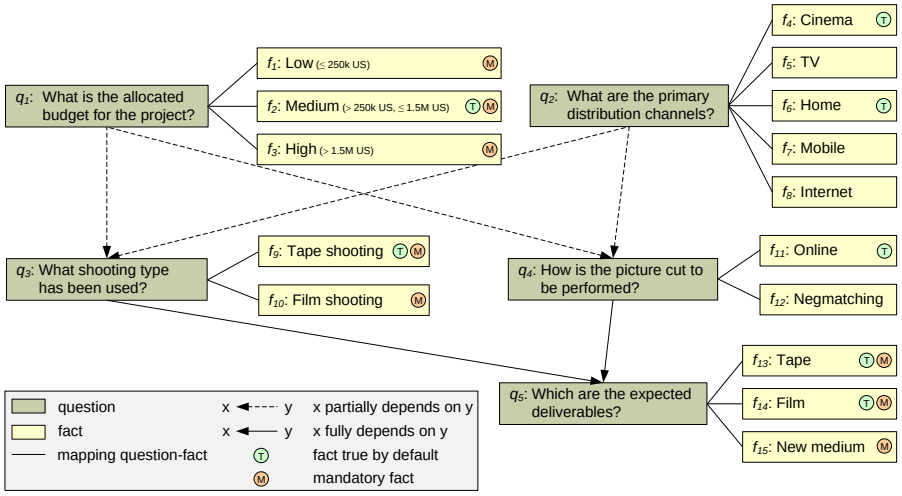

Fig. 17: A questionnaire model for post-production.

In Figure 17, there are questions that gather information on the type of shooting media (q_3), picture cut (q_4) and deliverables (q_5). These questions capture domain facts similar to the features in the feature diagram of Figure 15. Next to these questions, however, we have defined two high-level questions: question q_1, which enquires about the estimated budget for a post-production project (low, medium or high), and question q_2, which inquires about the distribution channel (e.g. cinema, TV, home). Different than feature diagrams, where each feature is mapped to a single alternative of a variation point in the customizable process model, "high-level" questions are defined with the intention of configuring multiple variation points at once, as shown later.

In general, one cannot freely answer questions because of interdependencies. For example, the answers to the questions of Figure 17 are interrelated as there is interplay among their facts due to the constraints imposed by post-production. In fact, negmatching (f_{12} in q_4) is a costly operation that can only be chosen if the project is shot on film (i.e. if f_{10} is true in q_3). However, the choice of which shooting medium to use is influenced by the project budget and by the distribution channel. For low budget productions (f_1 set to true in q_1), shooting and finishing on film are not allowed (hence the corresponding domain facts f_{10} and f_{14} must be set to false). In turn, if shooting on film is not allowed (f_{10} = false), negmatching must also be denied (f_{12} = false), and so on. This interplay among domain facts can be encoded by a set of *domain constraints* expressed as boolean formulae over the values of the domain facts, similar to the constraints defined among the features of a feature model. A *domain configuration* is thus a valuation of domain facts, resulting from answering a questionnaire, which does not violate the domain constraints.

A further difference between feature models and questionnaire models is the ability in a questionnaire model to establish an order for posing questions to users. This is achieved via *order dependencies*. A *partial* dependency captures an optional precedence between two questions: e.g. q_3 in Figure 17 can be posed after q_1 OR q_2 have been answered. A full dependency captures a mandatory precedence: e.g. q_5 is posed

after q_3 AND q_4. Dependencies can be set in a way to give priority to the most discriminating questions, i.e. the high-level questions q_1 and q_2 in our example, so that subsequent questions can be (partly) answered, automatically, by enforcing the domain constraints. If, for example, we pick "Low" budget in q_1, the questions about the shooting and cut type (q_3 and q_4) become irrelevant, because one can only choose facts "Tape shooting", respectively, "Online" editing, and will thus be skipped. These order dependencies can be arbitrary so long as cycles are avoided.

Questionnaire models offer both abstraction and guidance for the configuration of process models. Users can answer a questionnaire using an interactive questionnaire tool called *Quaestio*,[15] that poses questions in an order consistent with the order dependencies, and prevents users from entering conflicting answers to subsequent questions by dynamically enforcing the domain constraints. The tool also takes care of skipping questions that have become irrelevant while answering the questionnaire. Further guidance is provided in the form of recommendations attached to single questions and domain facts, providing contextual information on how to answer the questionnaire.

Questionnaire models have been applied to the configuration of C-iEPCs (cf. Section 6.1) and Configurable Workflows (cf. Section 6.2). The questionnaire model is linked to the customizable process model by assigning a *process fact* to each customization option of a configurable node in the customizable process model [La Rosa et al. 2008; La Rosa 2009]. A process fact is a boolean variable that captures the selection of a specific customization option of a configurable node in the customizable process model: a process fact set to true means that the corresponding option in the process model has been selected; vice versa, setting the fact to false means that the option is not selected. Different than feature models, there is not necessarily a one-on-one mapping between process facts and domain facts. Rather, a boolean expression over the domain facts of the questionnaire model is assigned to each process fact so that the latter is set to true when the corresponding expression evaluates to true. Thus, depending on how this mapping is defined, the customization of a single configurable node can be affected by multiple domain facts, as well as a single domain fact can affect the configuration of multiple configurable nodes. For example, we can map the questionnaire model of Figure 17 to the C-iEPC example of Figure 4 in such a way that when q_1 is answered with a "Low" budget level, all the configurable OR gateways in the process model get customized, at once, to their left-hand side flows. This is because a low budget production imposes that shooting, editing and release are all done on tape.

C.3. Decision Tables

Decision tables are an alternative, tabular representation of questionnaire models. A decision table is composed of *decisions* (also called *conditions*). A decision, which can be expressed in the form of a question, is associated with an enumerated set of possible *resolutions*. Each resolution can be linked to one or many variation points in a process model via an *effect* (also called *action*). The effect explains how the variation point needs to be customized when a particular resolution is taken.

Decision tables have been suggested as an abstraction mechanism in the KoBrA method (cf. Section B.0.1) and in the approach by Ripon et al. (cf. Section B.0.5). However, while decisions can be ordered in a decision table, the approaches that resort to this technique do not provide any mechanism to skip irrelevant decisions nor recommendations for taking the decisions.

[15] Quaestio is part of Synergia and of Apromore, see `http://processconfiguration.com` and `http://apromore.org`

D. TECHNIQUES FOR CORRECTNESS SUPPORT

Process model customization may lead to correctness issues in the customized model. If on the one hand, structural errors such as disconnected model elements are easy to detect and fix, on the other hand, behavioral anomalies such as deadlocks and livelocks pose challenges, as many such errors can only be identified via a state-space analysis of the model, which is exponential in complexity. A customizable process model capturing a realistic scenario can easily induce a large number of possible customizations. For example, if we assume 50 variation points each with three alternatives, we get $3^{50} \approx 7.18e + 23$ possible customizations. Checking the behavioral correctness of each individual customization *a-posteriori* is thus unfeasible.

In this section we discuss two techniques that can be used to guarantee both structural and behavioral correctness of customized process models *a-priori*, i.e. while customizing the process model. The techniques rely on the notion of hiding and blocking and have been applied to C-iEPCs (cf. Section 6.1) and Configurable Workflows (cf. Section 6.2). Since it has been shown that any behavioral restriction of process model behavior can be explained by applying the hiding and blocking operations to the activities of a process model [van der Aalst and Basten 2002], the correctness techniques presented in this section can in principle be adapted to offer correctness support to all approaches that customize process models by restriction.

D.1. Constraints Inference

The work in [van der Aalst et al. 2010] proposes a formal framework for transforming customizable process models incrementally, while preserving both structural and behavioral correctness. The framework is based on a technique to automatically infer propositional logic constraints from the control-flow dependencies of a process model (i.e. from its syntax), that, if satisfied by a customization step, guarantee the syntactic correctness of the customized model.

The theory was first developed in the context of Workflow nets and then extended to a subset of C-iEPCs. Workflow nets are a class of Petri nets specifically designed to model business processes. They come with a definition of *soundness* which ensures a process model to be free of behavioral anomalies such as deadlocks and lack of synchronization. Each Workflow nets activity (called *transition*) represents a process activity and can serve as a variation point: it is allowed by default and can be hidden or blocked during customization.

Whenever an activity is hidden or blocked, the current set of constraints is evaluated. If the constraints are satisfied, the customization step is applied. If the constraints are violated, a reduced propositional logic formula is computed, from which additional activities are identified that need to be customized simultaneously in order to preserve the syntactic correctness. For example, if an activity in the customizable process model is blocked, all nodes in a path starting with that activity need also to be removed to avoid disconnected elements. The set of constraints is incrementally updated after each step of the customization procedure.

A core result of this technique is that, for Workflow nets satisfying the *free-choice* property [Desel and Esparza 1995], if the model resulting from a customization step starting from a sound Workflow nets is a Workflow nets (i.e. it is structurally correct), then this latter Workflow nets is also sound. This means that for this class of nets, customization steps that preserve structural correctness also preserve behavioral correctness. Thus, via this technique, both structural and behavioral correctness of the customized process model are guaranteed at each configuration step.

This technique provides an efficient way of ensuring syntactical correctness during customization. However, it assumes that the customizable process model is already sound and free-choice. The latter does not represent a significant limitation since the large majority of constructs of languages such as BPMN, EPCs or BPEL can be mapped

to Workflow nets in this class [Lohmann et al. 2009]. On the other hand, imposing the customizable process model to be sound may hinder the applicability of this approach in practice. In fact, abstracting from data and resources may generate false positives (e.g., models that have behavioral problems due to data dependencies) and false negatives (e.g., the reported error is circumvented using data conditions).

D.2. Partner Synthesis

With the aim to address the shortcomings of [van der Aalst et al. 2010], the work in [van der Aalst et al. 2012] proposes a new technique for ensuring behavioral correctness during process configuration. This technique relies on the application of hiding and blocking to a wider Petri net class than free-choice Workflow nets, namely Open nets. Open nets can have multiple end states (whereas Workflow nets only have one) and can have complex non-free choice dependencies. Moreover, this technique relies on *weak termination* as a notion of behavioral correctness. Weak termination is a weaker correctness notion than soundness, as it only ensures that a process instance can always reach an end state from any state that can be reached from the start state. This means that even if some activities are unreachable (i.e. "dead"), the model will still be considered behaviorally correct. The authors argue that this correctness notion is more suitable for process model customization as parts of a customized model may be left intentionally dead.

The technique is based upon the notion of *configuration guideline*, which is inspired by the notion of operating guidelines used for partner synthesis [Wolf 2009]. A configuration guideline is a complete characterization of all *feasible* customizations, i.e. those customizations yielding a weakly terminating customized model (i.e. a correct model). These customizations are represented as possible combinations of blocked activities assuming that in the initial Open net all activities are allowed by default. Alternatively, it is possible to start from an Open net where all activities are blocked by default, and customize the net by allowing activities. In this case, the configuration guideline will store all possible combinations of allowed activities. Thus, one can check the configuration guideline at each customization step, and enforce further activities to be blocked or allowed, in order to keep the current customization feasible, in a way similar to the staged configuration approach in [van der Aalst et al. 2010]. However, the initial customizable Open net does not need to be sound. If the model has no feasible customizations, this is reported and the user will not be able to hide or block any activity.

The technique automatically generates the configuration guideline from an Open net. This is done by first building a *configuration interface* which can communicate with services that customize the original model. The configuration guideline thus represents the most permissible service that is able to interact with the Open net by customizing it.

This technique has been applied to the C-YAWL language, and implemented as a component of the YAWL Editor.[16] Once the tool has computed the configuration interface and its guideline for a C-YAWL model, the user can interactively customize the model. At each customization step, the system analyzes the configuration guideline and automatically blocks further YAWL ports to keep the current customization feasible. The user can also roll back a previously taken decision, e.g. by allowing a port that was blocked. In this case, the tool may impose that further ports have to be allowed in order to keep the configuration feasible. The tool's response time is instantaneous, because the traversal of the configuration guideline has linear complexity. The rationale of this approach is thus to move the computation time from customization-time to design-time, i.e. when the configuration guideline is built.

[16]See `http://yawlfoundation.org`

E. TERMINOLOGY

This appendix provides a list of all the relevant terms used in this paper and their definitions.

Activity variant: A *customization option* denoting a concrete refinement of an abstract activity, e.g. "Prepare tape for editing" is an activity variant of "Prepare medium for editing".

Adjustment point: A type of *variation point*; a sequence flow of the *customizable process model* which identifies the beginning or end of a fragment upon which a *change operation* can be applied. Adjustment points may be explicitly indicated in the model with a special notation, otherwise each flow is assumed to be an adjustment point.

Annotation: The graphical assignment of a *domain condition* to one or more *variation points*.

Change operation: The addition, deletion or modification of a fragment of the *customizable process model*.

Configurable model: See *configurable process model*.

Configurable node: A type of *variation point*; a node of the *customizable process model* whose behavior can be restricted. A configurable node may be an activity, an event, a gateway, or a resource or object associated with an activity.

Configurable process model: A *customizable process model* when *customization* is achieved by restriction of process behavior.

Customizable model: See *customizable process model*.

Customizable process model: Process model representing a family of *process model variants*, from which each variant can be derived via *transformations* after *customization decisions* are taken.

Customization: The process of deriving a *customized process model* from a *customizable process model* by applying *transformations* to *variation points*, based on *customization decisions* made by a user.

Customization decision: A choice made by a user in order to customize a customizable model. Customization decisions can be expressed at an abstract level in terms of *domain properties* (e.g. in post-production, one needs to choose the level of budget between high, medium and low), or at a concrete level in terms of *variation points* of the *customizable process model* itself (e.g. an OR-split can be restricted to an AND-split or to an XOR-split).

Customization option: A possible restriction or extension of the behavior of a *variation point*, e.g. a configurable activity has three customization options: on (the activity is intended to be kept), off (the activity is intended to be removed) and optional (the decision to remove the activity is deferred to run-time).

Customized model: See *customized process model*.

Customized process model: The process model obtained from the application of a *transformation algorithm* on a *customizable process model*; the result of *customization*. A customized process model represents a specific variant of the business process captured by the customizable process model, e.g. the variant of a claims handling process for home insurance.

Decision: See *customization decision*.

Decision table: A tabular representation of a *questionnaire model*.

Domain condition: A predicate over *domain properties*, e.g. "low budget = true".

Domain constraint: A restriction over the possible combinations of *domain properties*, e.g. "if low budget = true, then film shooting = false."

Domain model: A model of the variability of the domain in which the *customizable process model* is captured, defined as a set of *domain properties* and *domain constraints* over these properties.

Domain property: An entity of the domain of discourse, e.g. "low budget" or "film shooting".

Extension point: A *variation point* which can be customized by extension of process behavior.

Feature model: A type of *domain model* where *domain properties*, called features, are organized in a tree. Properties to be retained can be selected according to some *domain constraints*.

Process model variant: Process model capturing a variant of a specific business process, e.g. motor insurance and home insurance are two variants of a claims handling process. Process model variants can be consolidated in a single model called *customizable process model*.

Product hierarchy: A simplified *feature model* where *domain properties*, called products, are organized in a tree called product hierarchy, where *domain constraints* cannot be expressed.

Questionnaire model: A type of *domain model* where *domain properties*, called domain facts, are organized as a set of questions that need to be answered in order to choose which properties to retain, according to some *domain constraints*.

Transformation: A model change to customize a *variation point* of the *customizable process model* either by restricting or extending its behavior, e.g. a transformation may involve removing an activity altogether or adding a new one.

Transformation algorithm: The algorithm used to perform one or more *transformations*. A property of such an algorithm may be that of ensuring that the *customized process model* is syntactically correct.

Variability mechanism: A set of modeling constructs and their corresponding semantics to specify the relations between a *customizable process model* and its possible *customized process models*.

Variant: May refer to *activity variant* or *process model variant*.

Variation point: An element (i.e. a node or a sequence flow) of the *customizable process model* that can be customized via model *transformations*.

F. INTERCHANGEABILITY OF PROCESS MODELING LANGUAGES

We adopt the standard BPMN terminology when referring to process model elements, though we exemplify each approach using the approach's original process modeling language. This was done to preserve the original graphical notation provided by each approach. However, a process model that only features elements from the BPMN core set (sequence flow, activity, start and end events, XOR, AND and OR gateways), as in the examples used in this paper, can be converted into a model in any other language

(e.g. EPCs or UML ADs) and vice versa, using simple correspondences, as illustrated in Table XIII, barring some limitations (e.g. the OR-join is not supported by UML ADs).

EPCs deserve special consideration. This language imposes a strict alternation of functions and events, which can be intermitted by chains of connectors. Therefore, intermediate events (i.e. events that are neither start nor end events) must always be represented in EPCs, while they can be omitted in BPMN, as can intermediate conditions (representing states) in YAWL. In EPCs, an intermediate event captures a function's trigger or outcome, or the branching condition of an (X)OR-split. However, functions' triggers and outcomes may be omitted to maintain the alternation of functions and events, e.g. when an intermediate event is used to capture a branching condition, the preceding function does not require its outcome and the subsequent function does not require its trigger. As an example, Figure 18 shows a BPMN model and its corresponding EPC model, where intermediate events have been added to ensure the syntactical correctness of the model.

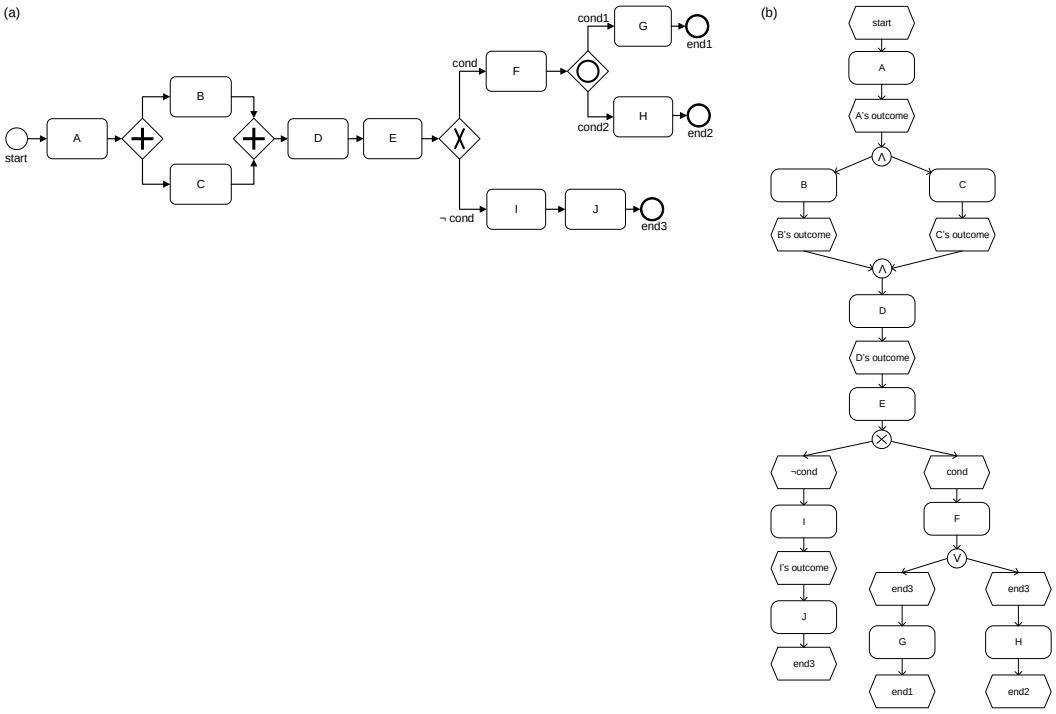

Fig. 18: An example BPMN model (a) and its corresponding model in EPCs (b).

More complex mappings are required when non-core BPMN elements such as boundary events are used. For example, [Dijkman et al. 2008] reports a rich mapping between a large subset of the BPMN language and Petri nets.

REFERENCES

ASADI, M., MOHABBATI, B., GRÖNER, G., AND GASEVIC, D. 2014. Development and validation of customized process models. *Journal of Systems and Software 96*, 73–92.

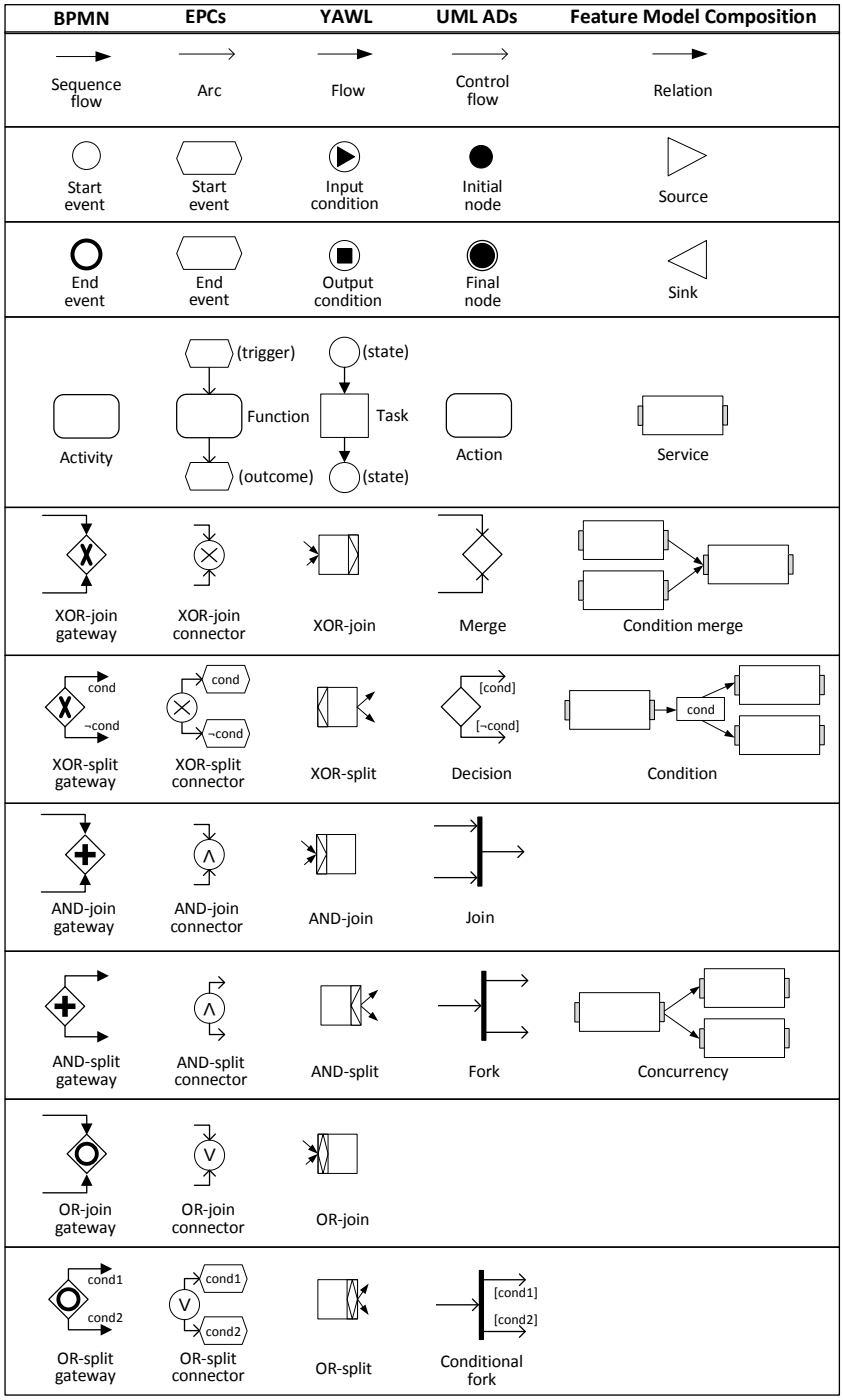

Table XIII: Mapping between BPMN core elements and corresponding constructs in EPCs, YAWL, UML ADs and Feature Model Composition.

ATKINSON, C., BAYER, J., AND MUTHIG, D. 2000. Component-Based Product Line Development: The KobrA Approach. In *Proceedings of the 1st Software Product Lines Conference (SPLC'00)*, P. Donohoe, Ed. Kluwer, 289–309.

ATKINSON, C., BOSTAN, P., BRENNER, D., FALCONE, G., GUTHEIL, M., HUMMEL, O., JUHASZ, M., AND STOLL, D. 2008. Modeling components and component-based systems in kobra. In *The Common Component Modeling Example: Comparing Software Component Models*. Lecture Notes in Computer Science Series, vol. 5153. Springer, 54–84.

AYORA, C., TORRES, V., WEBER, B., REICHERT, M., AND PALECHANO, V. 2014. VIVACE: A framework for the systematic evaluation of variability support in process-aware information systems. *Information and Software Technology*.

BATORY, D. 2005. Feature Models, Grammars, and Propositional Formulas. In *Proceedings of the 9th International Conference on Software Product Lines (SPLC'05)*, J. Obbink and K. Pohl, Eds. Lecture Notes in Computer Science Series, vol. 3714. Springer, 7–20.

CIUKSYS, D. AND CAPLINSKAS, A. 2007. Reusing ontological knowledge about business processes in is engineering: Process configuration problem. *Informatica, Lith. Acad. Sci. 18,* 4, 585–602.

CZARNECKI, K. AND EISENECKER, U. 2000. *Generative Programming: Methods, Tools, and Applications*. Addison-Wesley.

CZARNECKI, K., HELSEN, S., AND EISENECKER, U. 2005. Formalizing Cardinality-Based Feature Models and Their Specialization. *Software Process: Improvement and Practice 10,* 1, 7–29.

DESEL, J. AND ESPARZA, J. 1995. *Free Choice Petri Nets*. Cambridge Tracts in Theoretical Computer Science Series, vol. 40. Cambridge University Press.

DIJKMAN, R.M., DUMAS, M., AND OUYANG, C. 2008. Semantics and analysis of business process models in BPMN. *Information & Software Technology 50,* 12, 1281–1294.

DÖHRING, M., REIJERS, H., AND SMIRNOV, S. 2014. Configuration vs. adaptation for business process variant maintenance: an empirical study. *Inf. Syst. 39*, 108–133.

DÖHRING, M., ZIMMERMANN, B., AND KARG, L. 2011. Flexible Workflows at Design- and Runtime Using BPMN2 Adaptation Patterns. In *Proceedings of Business Information Systems*. Lecture Notes in Business Information Processing Series, vol. 87. Springer, 25–36.

FINK, A. 2010. *Conducting research literature reviews: from the internet to paper* 3rd edition Ed. Sage Publications.

GRÖNER, G., BOSKOVIC, M., PARREIRAS, F. S., AND GASEVIC, D. 2013. Modeling and validation of business process families. *Inf. Syst. 38,* 5, 709–726.

HEUER, A., STRICKER, V., BUDNIK, C., KONRAD, S., AND LAUENROTH, K. 2013. Defining variability in activity diagrams and Petri nets. *Science of Computer Programming 78*, 2414?2432.

KANG, K., COHEN, S., HESS, J., NOVAK, W., AND PETERSON, A. 1990. Feature-Oriented Domain Analysis (FODA) Feasibility Study. Technical Report CMU/SEI-90-TR-21, Software Engineering Institute, Carnegie Mellon University.

KITCHENHAM, B. 2004. Procedures for undertaking systematic reviews. Tech. rep., Computer Science Department, Keele University (TR/SE- 0401) and National ICT Australia Ltd. (0400011T.1).

KORHERR, B. AND LIST, B. 2007. A UML 2 profile for variability models and their dependency to business processes. In *Proceedings of the 18th International Workshops on Database and Expert Systems Applications, Regensburg, Germany*. IEEE Computer Society, 829–834.

KULKARNI, V. AND BARAT, S. 2011. Business process families using model-driven techniques. In *Proceedings of the Business Process Management Workshops and Education Track, Hoboken, NJ, USA*. Springer, 314–325.

LA ROSA, M. 2009. Managing variability in process-aware information systems. Phd thesis, Queensland University of Technology.

LA ROSA, M., GOTTSCHALK, F., DUMAS, M., AND VAN DER AALST, W. 2008. Linking Domain Models and Process Models for Reference Model Configuration. In *Business Process Management 2007 Workshops*, A. ter Hofstede, B. Benatallah, and H. Paik, Eds. Lecture Notes in Computer Science Series, vol. 4928. Springer, 417–430.

LA ROSA, M., VAN DER AALST, W., DUMAS, M., AND TER HOFSTEDE, A. 2009. Questionnaire-based Variability Modeling for System Configuration. *Software and Systems Modeling 8,* 2, 251–274.

LEVY, Y. AND ELLIS, T. 2006. A systems approach to conduct an effective literature review in support of information systems research. *Informing Science Journal 9*, 181–212.

LOHMANN, N., VERBEEK, E., AND DIJKMAN, R. 2009. Petri net transformations for business processes - a survey. *T. Petri Nets and Other Models of Concurrency 2*, 46–63.

LU, R., SADIQ, S., AND GOVERNATORI, G. 2009. "on managing business processes variants". *Data Knowl. Eng. 68,* 7, 642–664.

MACHADO, I., BONIFACIO, R., ALVES, V., TURNES, L., AND MACHADO, G. 2011. Managing Variability in Business processes: An Aspect-Oriented Approach. In *Proceedings of the Int. Workshop on Early Aspects*. ACM, 25–30.

MEERKAMM, S. 2010. Configuration of multi-perspectives variants. In *Business Process Management Workshops - BPM 2010 International Workshops and Education Track, Hoboken, NJ, USA, September 13-15, 2010, Revised Selected Papers*. 277–288.

NGUYEN, T., COLMAN, A. W., AND HAN, J. 2011. Modeling and managing variability in process-based service compositions. In *Proceedings of the 9th International Conference on Service-Oriented Computing (ICSOC), Paphos, Cyprus*. Springer, 404–420.

OKOLI, C. AND SCHABRAM, K. 2010. A guide to conducting a systematic literature review of information systems research. *Sprouts: Working Papers on Information Systems 10,* 26, 1–49.

RANDOLPH, J. 2009. A guide to writing the dissertation literature review. *Practical Assessment, Research & Evaluation 14,* 13, 1–13.

RASTREPKINA, M. 2010. Managing variability in process models by structural decomposition. In *Business Process Modeling Notation - Second International Workshop, BPMN 2010, Potsdam, Germany, October 13-14, 2010. Proceedings*. 106–113.

RAZAVIAN, M. AND KHOSRAVI, R. 2008. Modeling Variability in Business Process Models Using UML. In *Proceedings of the 5th International Conference on Information Technology: New Generations (ITGN'08)*, S. Latifi, Ed. 82–87.

RIPON, S., TALUKDER, K., AND MOLLA, K. 2010. Modelling variability for system families. *Malaysian Journal of Computer Science 16,* 1, 37–46.

ROWLEY, J. AND SLACK, F. 2004. Conducting a literature review. *Management Research News 27,* 6, 31–39.

SANTOS, E., PIMENTEL, J., CASTRO, J., SÁNCHEZ, J., AND PASTOR, O. 2010. Configuring the Variability of Business Process Models Using Non-Functional Requirements. In *Proceedings of EMMSAD*. Lecture Notes in Business Information Processing Series, vol. 50. Springer, 274–286.

SCHOBBENS, P.-Y., HEYMANS, P., AND TRIGAUX, J.-C. 2006. Feature Diagrams: A Survey and a Formal Semantics. In *Proceedings of the 14th IEEE International Conference on Requirements Engineering (RE'06)*, M. Glinz and R. Lutz, Eds. IEEE Computer Society, 136–145.

SCHUNSELAAR, D., VERBEEK, E., VAN DER AALST, W., AND REIJERS, H. 2011. Creating sound and reversible configurable process models using CoSeNets. Tech. Rep. BPM-11-21, BPM Center Report.

SCHUNSELAAR, D., VERBEEK, E., VAN DER AALST, W., AND REIJERS, H. 2012. Creating sound and reversible configurable process models using CoSeNets. In *Proc. of Business Information Systems*. Lecture Notes in Business Information Processing Series, vol. 117. Springer, 24–35.

TORRACO, R. 2005. Writing integrative literature reviews: guidelines and examples. *Human Resource Development Review 4,* 3, 356–367.

VALENCA, G., ALVES, C., ALVES, V., AND NIU, N. 2013. A systematic mapping study on business process variability. *Int. Journal of Computer Science & Information Technology 5,* 1, 1–21.

VAN DER AALST, W. AND BASTEN, T. 2002. Inheritance of Workflows: An Approach to Tackling Problems Related to Change. *Theoretical Computer Science 270,* 1-2, 125–203.

VAN DER AALST, W., DUMAS, M., GOTTSCHALK, F., TER HOFSTEDE, A., LA ROSA, M., AND MENDLING, J. 2010. Preserving Correctness During Business Process Model Configuration. *Formal Aspects of Computing 22,* 3, 459–482.

VAN DER AALST, W., LOHMANN, N., AND LA ROSA, M. 2012. Correctness Ensuring Process Configuration: An Approach Based on Partner Synthesis. *Information Systems 37,* 6, 574–592.

WEBSTER, J. AND WATSON, R. 2002. Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly 26,* 2, xiii–xxiii.

WOLF, K. 2009. Does my Service Have Partners? In *Transactions on Petri Nets and Other Models of Concurrency II*, K. Jensen and W. van der Aalst, Eds. Lecture Notes in Computer Science Series, vol. 5460. Springer-Verlag, Berlin, 152–171.