

Correctness Ensuring Process Configuration: An Approach Based on Partner Synthesis (extended version)

Wil van der Aalst^{1,3}, Niels Lohmann^{1,2}, Marcello La Rosa³, and Jingxin Xu³

¹ Eindhoven University of Technology, The Netherlands

w.m.p.v.d.aalst@tue.nl

² Universität Rostock, Germany

niels.lohmann@uni-rostock.de

³ Queensland University of Technology, Australia

m.larosa@qut.edu.au, jingxin.xu@connect.qut.edu.au

Abstract. A configurable process model describes a family of similar process models in a given domain. Such a model can be configured to obtain a *specific* process model that is subsequently used to handle individual cases, for instance, to process customer orders. *Process configuration is notoriously difficult as there may be all kinds of interdependencies between configuration decisions.* In fact, an incorrect configuration may lead to behavioral issues such as deadlocks and livelocks. To address this problem, we present a new verification approach inspired by the “operating guidelines” used for partner synthesis. We view the configuration process as an external service, and compute a characterization of all such services which meet particular requirements via the notion of *configuration guideline*. As a result, we can characterize all feasible configurations (i. e., configurations without behavioral problems) at design time, instead of repeatedly checking each individual configuration while configuring a process model.

Key words: Configurable process model, operating guideline, Petri net

1 Introduction and Background

Although large organizations support their processes using a wide variety of process-aware information systems, the majority of business processes are still not directly driven by process models. Despite the success of Business Process Management (BPM) thinking in organizations, Workflow Management (WfM) systems — today often referred to as *BPM systems* — are not widely used. One of the main problems of BPM technology is the “lack of content”; that is, providing just a generic infrastructure to build process-aware information systems is insufficient as organizations need to support specific processes. Organizations want to have “out-of-the-box” support for standard processes and are only willing to design and develop system support for organization-specific processes. Yet most BPM systems expect users to model basic processes from scratch.

Enterprise Resource Planning (ERP) systems such as SAP and Oracle, on the other hand, focus on the support of these common processes. Although all ERP systems have workflow engines comparable to the engines of BPM systems, the majority of processes are not supported by software which is driven by models. For example, most of SAP’s functionality is not grounded in their workflow component, but hard-coded in application software. ERP vendors try to capture “best practices” in dedicated applications designed for a particular purpose. Such systems can be configured by setting parameters. System configuration can be a time consuming and complex process. Moreover, configuration parameters are exposed as “switches in the application software”, thus making it difficult to see the impact of certain settings and dependencies between them.

A model-driven process-oriented approach toward supporting business processes has all kinds of benefits ranging from improved analysis possibilities (verification, simulation, etc.) and better insights, to maintainability and ability to rapidly develop organization-specific solutions. Although obvious, this approach has not been adopted thus far, because BPM vendors failed to provide content and ERP vendors suffered from the “Law of the handicap of a head start”. ERP vendors managed to effectively build data-centric solutions to support particular tasks. However, the complexity and large installed base of their products makes it impossible to refactor their software and make it process-centric.

Based on the limitations of existing BPM and ERP systems, we propose to use *configurable process models*. A configurable process model represents a *family of process models*; that is, a model that through configuration can be customized for a particular setting. Configuration is achieved by *hiding* (i. e., bypassing) or *blocking* (i. e., inhibiting) certain fragments of the configurable process model [12]. This way, the desired behavior is selected. From the viewpoint of generic BPM software, configurable process models can be seen as a mechanism to add content to these systems. By developing comprehensive collections of configurable models, particular domains can be supported. From the viewpoint of ERP software, configurable process models can be seen as a means to make these systems more process-centric, although in the latter case, quite some refactoring is needed as processes are hidden in table structures and application code.

Various configurable languages have been proposed as extensions of existing languages (e. g., C-EPCs [21], C-iEPCs [15], C-WF-nets [3], C-SAP, C-BPEL) but few are actually supported by enactment software (e. g., C-YAWL [13]). In this paper, we are interested in models in the latter class of languages, which, unlike traditional reference models [8,7,11], are executable after they have been configured. Specifically, we focus on the *verification of configurable executable process models*. In fact, because of hiding and/or blocking selected fragments, the instances of a configured model may suffer from behavioral anomalies such as deadlocks and livelocks. This problem is exacerbated by the total number of possible configurations a model may have, and by the complex dependencies which may exist between various configuration options. For example, the configurable process model we constructed from the VICS documentation⁴ — an

⁴ See www.vics.com (Voluntary Interindustry Commerce Solutions Association).

industry standard for logistics and supply chain management—comprises 50 activities. Each of these activities may be “blocked”, “hidden”, or “allowed”, depending on the configuration requirements. This results in $3^{50} \approx 7.18e+23$ possible configurations. Clearly, checking the *feasibility* of each single configuration can be time consuming as this would typically require performing state-space analysis. Moreover, characterizing the “family of correct models” for a particular configurable process model is even more difficult and time-consuming as a naive approach would require solving an exponential number of state-space problems.

As far as we know, our earlier approach described in [3] is the only one focusing on the verification of configurable process models. Other approaches discuss syntactical correctness related to configuration [21,9,7], but do not provide techniques for ensuring the behavioral correctness of the configured models. In this paper, we propose a completely novel verification approach where we consider the configuration process as an “external service” and then synthesize a “most permissive partner” using the approach described by Wolf [22] and implemented in the tool Wendy [19]. This most permissive partner is closely linked to the notion of *operating guidelines* for service behavior [18]. In this paper, we define for any configurable model a so-called *configuration guideline* to characterize all correct process configurations. This approach provides the following advantages over our previous approach [3]:

- We provide a *complete characterization of all possible configurations at design time*; that is, the *configuration guideline*.
- Computation time is moved *from configuration time to design time* and results can be reused more easily.
- *No restrictions are put on the class of models* which can be analyzed. The previous approach [3] was limited to sound free-choice WF-nets. Our new approach can be applied to models which do not need to be sound, which can have complex (non-free choice) dependencies, and which can have multiple end states.

To prove the practical feasibility of this new approach, we have implemented it as a plugin of the toolset supporting C-YAWL [23].

The remainder of this paper is organized as follows. Section 2 introduces basic concepts such as open nets and weak termination. These concepts are used in Section 3 to formalize the notion of process configuration. Section 4 presents the solution approach for correctness ensuring configuration. Section 5 discusses tool support, and Sect. 6 concludes the paper.

2 Business Process Models

For the formalization of the problem we use Petri nets, which offer a formal model of concurrent systems. However, the same ideas can be applied to other languages (e. g. C-YAWL, C-BPEL), as it is easy to map the core structures of these languages onto Petri nets. Moreover, our analysis approach is quite generic and does not rely on specific Petri net properties.

Definition 1 (Petri net). A marked Petri net is a tuple $N = (P, T, F, m_0)$ such that: P and T ($P \cap T = \emptyset$) are finite sets of places and transitions, respectively, $F \subseteq (P \times T) \cup (T \times P)$ is a flow relation, and $m_0 : P \rightarrow \mathbb{N}$ is an initial marking.

A Petri net is a directed graph with two types of nodes: places and transitions, which are be connected by arcs as specified in the flow relation. If $p \in P$, $t \in T$, and $(p, t) \in F$, then place p is an input place of t . Similarly, $(t, p) \in F$ means that p is an output place of t .

The *marking* of a Petri net describes the distribution of tokens over places and is represented by a *multiset of places*. For example, the marking $m = [a^2, b, c^4]$ indicates that there are two tokens in place a , one token in b , and four tokens in c . Formally m is a function such that $m(a) = 2$, $m(b) = 1$, and $m(c) = 4$. We use \oplus to compose multisets; for instance, $[a^2, b, c^4] \oplus [a^2, b, d^2, e] = [a^4, b^2, c^4, d^2, e]$.

A transition is *enabled* and can *fire* if all its input places contain at least one token. Firing is atomic and consumes one token from each of the input places and produces one token on each of the output places. $m_0 \xrightarrow{t} m$ means that t is enabled in marking m_0 and the firing of t in m_0 results in marking m . We use $m_0 \xrightarrow{*} m$ to denote that m is reachable from m_0 ; that is, there exists a (possibly empty) sequence of enabled transitions leading from m_0 to m .

For our configuration approach, we use *open nets*. Open nets extend classical Petri nets with the identification of final markings Ω and a labeling function ℓ .

Definition 2 (Open net). A tuple $N = (P, T, F, m_0, \Omega, L, \ell)$ is an open net if

- (P, T, F, m_0) is a marked Petri net (called the inner net of N),
- $\Omega \subset P \rightarrow \mathbb{N}$ is a finite set of final markings,
- L is a finite set of labels,
- $\tau \notin L$ is a label representing invisible (also called silent) steps, and
- $\ell : T \rightarrow L \cup \{\tau\}$ is a labeling function.

We use *transition labels* to represent the activity corresponding to the execution of a particular transition. Moreover, if an activity appears multiple times in a model, we use the same label to identify all the occurrences of that activity. The special label τ refers to an invisible step, sometimes referred to as “silent”. Invisible transitions are typically use to represent internal actions which do not mean anything at the business level, cf. the “inheritance of dynamic behavior” framework [2,6]. In Section 4 we use visible labels to synchronize two open nets. However, initially we use labels to denote activities that may be configured.

Figure 1 shows an example open net which models a typical travel request approval. The process starts with the preparation of the travel form. This can either be done by the employee or be delegated to a secretary. In both cases, the employee personally needs to arrange the travel insurance. If the form has been prepared by the secretary, the employee needs to check it before submitting it for approval. The administrator can then approve or reject the request, or make a request for change. Now, the employee can update the form according to the administrator’s suggestions and resubmit it. In Fig. 1, all transitions bear a

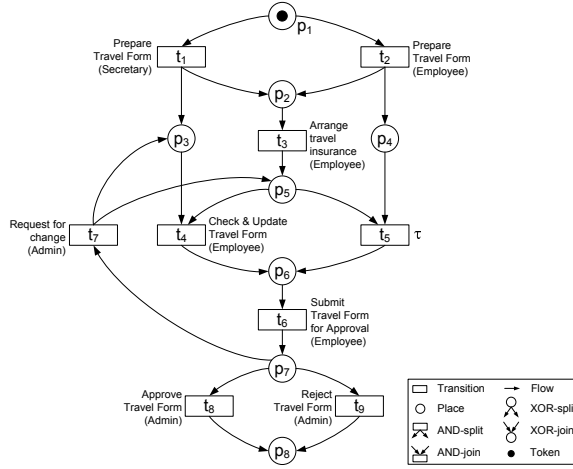


Fig. 1. The open net for travel request approval ($\Omega = \{[p_8]\}$).

unique label, except for t_5 which bears a τ -label as it has only been added for routing purposes.

Unlike our previous approach [3] based on WF-nets [1] and hence limited to a single final place, here we allow *multiple final markings*. Good runs of an open net end in a marking in set Ω . Therefore, an open net is considered to be erroneous if it can reach a marking from which no final marking can be reached any more. An open net *weakly terminates* if a final marking is reachable from every reachable marking.

Definition 3 (Weak termination). *An open net $N = (P, T, F, m_0, \Omega, L, \ell)$ weakly terminates if and only if (iff) for any marking m with $m_0 \xrightarrow{*} m$ there exists a final marking $m_f \in \Omega$ such that $m \xrightarrow{*} m_f$.*

The net in Fig. 1 is weakly terminating. Weak termination is a weaker notion than soundness, as it does not require every transition to be quasi-live [1]. This correctness notion is more suitable as parts of a correctly configured net may be left dead intentionally.

3 Process Model Configuration

We use open nets to model configurable process models. An open net can be configured by blocking or hiding transitions which bear a visible label (i. e., not a τ -label). Blocking a transition means that the corresponding activity is no longer available and none of the paths with that transition cannot be taken any more. Hiding a transition means that the corresponding activity is bypassed, but paths with that transition can still be taken. If a transition is neither blocked nor hidden, we say it is allowed, meaning nothing changes in the model. Configuration is achieved by setting visible labels to *allow*, *hide* or *block*.

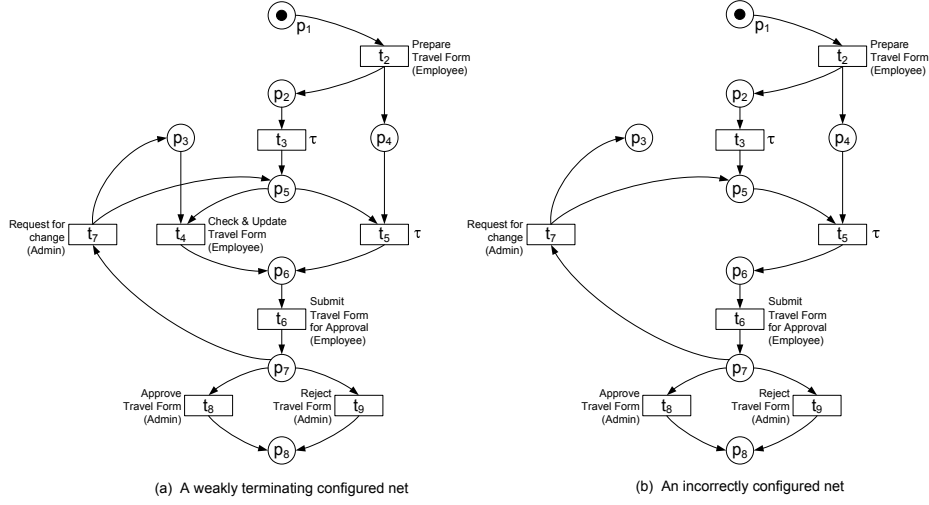


Fig. 2. Two possible configured nets based on the model in Fig. 1.

Definition 4 (Open net configuration). Let N be an open net with label set L . A mapping $C_N : L \rightarrow \{allow, hide, block\}$ is a configuration for N . We define:

- $A_N^C = \{t \in T \mid \ell(t) \neq \tau \wedge C_N(\ell(t)) = allow\}$,
- $H_N^C = \{t \in T \mid \ell(t) = \tau \vee C_N(\ell(t)) = hide\}$, and
- $B_N^C = \{t \in T \mid \ell(t) \neq \tau \wedge C_N(\ell(t)) = block\}$.

An open net configuration implicitly defines an open net, called *configured net*, where the blocked transitions are removed and the hidden transitions are given a τ -label.

Definition 5 (Configured net). Let $N = (P, T, F, m_0, \Omega, L, \ell)$ be an open net and C_N a configuration of N . The resulting configured net $\beta_N^C = (P, T^C, F^C, m_0, \Omega, L, \ell^C)$ is defined as follows:

- $T^C = T \setminus (B_N^C)$,
- $F^C = F \cap ((P \cup T_C) \times (P \cup T_C))$, and
- $\ell^C(t) = \ell(t)$ for $t \in A_N^C$ and $\ell^C(t) = \tau$ for $t \in H_N^C$.

As an example, Fig. 2(a) shows the configured net derived from the open net in Fig. 1 and the configuration $C_N(Prepare\ Travel\ Form\ (Secretary)) = block$ (to allow only employees to prepare travel forms), $C_N(Arrange\ Travel\ Insurance\ (Employee)) = hide$ (to skip arranging the travel insurance), and $C_N(x) = allow$ for all other labels x .

Typically, configurable process models cannot be freely configured, because the application of hiding and blocking has to comply with the application domain in which the model has been constructed. For instance, in the travel request example we cannot hide the labels of both t_1 and t_2 , because all the other

activities depend on the preparation of the travel form, nor block the label of t_8 , because there must be an option to approve the request. The link between configurable process models and domain decisions was explored in [16].

A configured net may have disconnected nodes and some parts may be dead (i. e., can never become active). Such parts can easily be removed. However, as we impose no requirements on the structure of configurable models, these disconnected or dead parts are irrelevant with respect to weak termination. For example, if we block the label of t_2 in Fig. 1, transition t_5 becomes dead as it cannot be enabled any more, and hence can also be removed without causing any behavioral issues. Nonetheless, not every configuration of an open net results in a weakly terminating configured net. For example, by blocking the label of t_4 in the configured net of Fig. 2(a), we obtain the configured net in Fig. 2(b). This net is not weakly terminating because after firing t_7 tokens will get stuck in p_3 (as this place does not have any successor) and in p_5 (as t_5 can no longer fire).

Blocking can cause behavioral anomalies such as the deadlock in Fig. 2(b). However, hiding cannot cause such issues, because it merely changes the labels of an open net. Hence, we shall focus on blocking rather than hiding.

In light of the above, in this paper we are interested in all configurations which yield weakly terminating configured nets. We use the term *feasibility* to refer to such configured nets.

Definition 6 (Feasible configuration). *Let N be an open net and C_N a configuration of N . C_N is feasible iff the configured net β_N^C weakly terminates.*

More precisely, given a configurable process model N , we are interested in the following two questions:

- Is a *particular* configuration C_N feasible?
- How to characterize the set of *all* feasible configurations?

The remainder of this paper is devoted to a new verification approach answering these questions. This approach extends the work in [3] in two directions: (i) it imposes *no unnecessary requirements* on the configurable process model (allowing for non-free-choice nets [10] and nets with multiple end places/markings), and (ii) it checks a *weaker correctness* notion (i. e. weak termination instead of soundness). For instance, the net in Fig. 1 is not free-choice because t_4 and t_5 share an input place, but their sets of input places are not identical. The non-free-choice construct is needed to model that after firing t_1 or t_7 , t_5 cannot be fired, and similarly, after firing t_2 , t_4 cannot be fired.

4 Correctness Ensuring Configuration

To address the two main questions posed in the previous section, we could use a direct approach by enumerating all possible configurations and simply checking whether each of the configured nets β_N^C weakly terminates or not. As indicated before, the number of possible configurations is exponential in

the number of configurable activities. Moreover, most techniques for checking weak termination typically require the construction of the state space. Hence, traditional approaches are computationally expensive and do not yield a useful characterization of the set of all feasible configuration. Consequently, we propose a completely different approach using the synthesis technique described in [22]. *The core idea is to see the configuration as an “external service” and then synthesize a “most permissive partner”.* This most permissive partner represents all possible “external configuration services” which yield a feasible configuration. The idea is closely linked to the notion of *operating guidelines* for service behavior [18]. An operating guideline is a finite representation of all possible partners. Similarly, our *configuration guideline characterizes all feasible process configurations*. This configuration guideline can also be used to *efficiently check the feasibility of a particular configuration without exploring the state space of the configured net*. Our approach consists of two steps:

1. Transform the configurable process model (represented as an open net N) into a *configuration interface* N^{CI} .
2. Synthesize the “most permissive partner” for the configuration interface constructed in N^{CI} . This is the configuration guideline for N .

For our solution approach, we compose the configurable process model with a “configuration service”. To do so, we first introduce *composition*. Open nets can be composed by synchronizing transitions according to their visible labels. In the resulting net, all transitions bear a τ -label and labeled transitions without counterpart in the other net disappear.

Definition 7 (Composition). *For $i \in \{1, 2\}$, let $N_i = (P_i, T_i, F_i, m_{0_i}, \Omega_i, L_i, \ell_i)$ be open nets. N_1 and N_2 are composable iff the inner nets of N_1 and N_2 are pairwise disjoint. The composition of two composable open nets is the open net $N_1 \oplus N_2 = (P, T, F, m_0, \Omega, L, \ell)$ with:*

- $P = P_1 \cup P_2$,
- $T = \{t \in T_1 \cup T_2 \mid \ell(t) = \tau\} \cup \{(t_1, t_2) \in T_1 \times T_2 \mid \ell(t_1) = \ell(t_2) \neq \tau\}$,
- $F = (F_1 \cup F_2) \cap ((P \times T) \cup (T \times P)) \cup \{(p, (t_1, t_2)) \in P \times T \mid (p, t_1) \in F_1 \vee (p, t_2) \in F_2\} \cup \{((t_1, t_2), p) \in T \times P \mid (t_1, p) \in F_1 \vee (t_2, p) \in F_2\}$,
- $m_0 = m_{0_1} \oplus m_{0_2}$,
- $\Omega = \{m_1 \oplus m_2 \mid m_1 \in \Omega_1 \wedge m_2 \in \Omega_2\}$,
- $L = \emptyset$, and $\ell(t) = \tau$ for $t \in T$.

Via composition, the behavior of each original net can be limited; for instance, transitions may no longer be available or may be blocked by one of the two original nets. Hence, it is possible that N_1 and N_2 are weakly terminating, but $N_1 \oplus N_2$ is not. Similarly, $N_1 \oplus N_2$ may be weakly terminating, but N_1 and N_2 are not. The labels of the two open nets in Def. 7 serve now a different purpose: they are not used for configuration, but for synchronous communication as described in [22].

With the notions of composition and weak termination, we define the *controllability*. We need this concept to reason about the existence of feasible configurations.

Definition 8 (Controllability). An open net N is controllable iff there exists an open net N' such that $N \oplus N'$ is weakly terminating.

Open net N' is called a *partner* of N if $N \oplus N'$ is weakly terminating. Hence, N is controllable if there exists a partner. Wolf [22] presents an algorithm to check controllability: if an open net is controllable, this algorithm can synthesize a partner.

After these preliminaries, we define the notion of a *configuration interface*. One of the objectives of this paper was to characterize the set of all feasible configurations by synthesizing a “most permissive partner”. To do this, we transform a configurable process model (i. e., an open net N) into an open net N^{CI} , called the configuration interface, which can communicate with services which configure the original model. In fact, we shall provide two configuration interfaces: one where everything is *allowed by default* and the external configuration service can block labels, and the other where everything is *blocked by default* and the external configuration service can allow labels. In either case, the resulting open net N^{CI} is controllable iff there exists a feasible configuration C_N of N . Without loss of generality, we assume a 1-safe initial marking; that is, $m_0(p) > 0$ implies $m_0(p) = 1$. This assumption helps simplifying the configuration interface.

Definition 9 (Configuration interface; allow by default). Let $N = (P, T, F, m_0, \Omega, L, \ell)$ be an open net. We define the open net with configuration interface $N_a^{CI} = (P^C, T^C, F^C, m_0^C, \Omega^C, L^C, \ell^C)$ with

- $T^V = \{t \in T \mid \ell(t) \neq \tau\}$,
- $P^C = P \cup \{p_{start}\} \cup \{p_t^a \mid t \in T^V\}$,
- $T^C = T \cup \{t_{start}\} \cup \{b_x \mid x \in L\}$,
- $F^C = F \cup \{(p_{start}, t_{start})\} \cup \{(t_{start}, p) \mid p \in P \wedge m_0(p) = 1\} \cup \{(t, p_t^a) \mid t \in T^V\} \cup \{(p_t^a, t) \mid t \in T^V\} \cup \{(b_x, p_{start}) \mid x \in L\} \cup \{(p_{start}, b_x) \mid x \in L\} \cup \{(p_t^a, b_{\ell(t)}) \mid t \in T^V\}$,
- $m_0^C = [p^1 \mid p \in \{p_{start}\} \cup \{p_t^a \mid t \in T^V\}]$,
- $\Omega^C = \{m \oplus \bigoplus_{t \in T} m_t^* \mid m \in \Omega \wedge \forall_{t \in T} m_t^* \in \{[], [p_t^a]\}\}$,
- $L^C = \{start\} \cup \{block_x \mid x \in L\}$
- $\ell^C(t_{start}) = start$, $\ell^C(b_x) = block_x$ for $x \in L$, and $\ell^C(t) = \tau$ for $t \in T$.

Figure 3 illustrates the two configuration interfaces for a simple open net N . In both interfaces, the original net N consisting of places $\{p_1, p_2, p_3, p_4\}$ and transitions $\{t_1, t_2, t_3, t_4\}$ is retained, but all transition labels are set to τ . Let us focus on the configuration interface where all activities are allowed by default (Fig. 3(b)). Here transitions b_x and b_y are added to model the blocking of labels x and y , respectively. Places $p_{t_1}^a$, $p_{t_2}^a$, and $p_{t_3}^a$ are also added to connect the new transitions to the existing ones, and are initially marked as all configurable transitions are allowed by default. Firing b_x will block t_1 and t_2 by removing the tokens from $p_{t_1}^a$ and $p_{t_2}^a$. These two transitions are blocked at the same time because both bear the same label x in N . Firing b_y will block t_3 . Transitions b_x and b_y are labeled respectively $block_x$ and $block_y$. This means that in the composition with a partner they can only fire if a corresponding transition in the

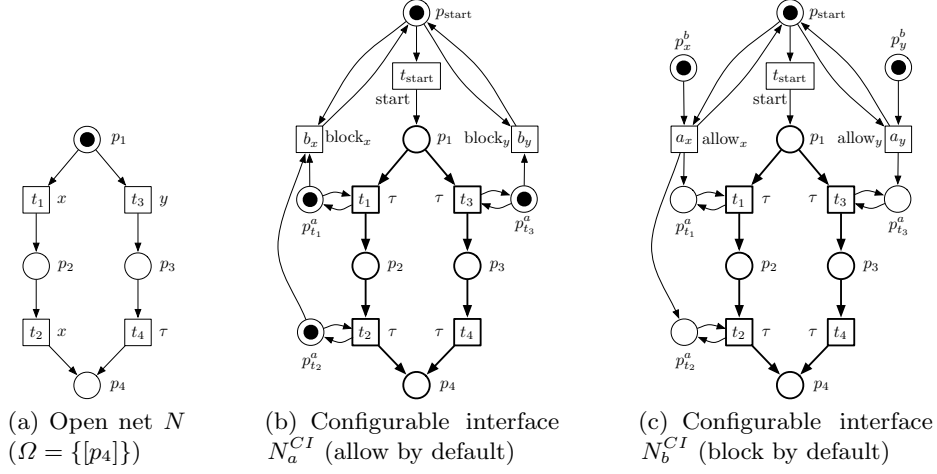


Fig. 3. An example open net (a) and its two configuration interfaces (b,c).

partner can fire. Transition $start$ has been added to ensure configuration actions take place *before* the original net is activated. This way, we avoid “configuration on the fly”. We shall discuss the construction of the configuration interface where all activities are blocked by default later on.

Consider now a configuration service represented as an open net Q . $N_a^{CI} \oplus Q$ is the composition of the original open net (N) extended with a configuration interface (N_a^{CI}), and the configuration service Q . First, blocking transitions such as b_x and b_y can fire (apart from unlabeled transitions in Q). Next, transition $start$ fires after which blocking transitions such as b_x and b_y can no longer fire. Hence, only the original transitions in N_a^{CI} can fire in the composition. The configuration service Q may still execute transitions, but these cannot influence N_a^{CI} any more. Hence, Q represents a feasible configuration iff N_a^{CI} can reach one of its final markings from any reachable marking in the composition. So Q corresponds to a feasible configuration iff $N_a^{CI} \oplus Q$ is weakly terminating; that is, Q is a partner of N_a^{CI} .

To illustrate the basic idea, we introduce the notion of a *canonical configuration partner*; that is, the representation of a configuration $C_N : L \rightarrow \{allow, hide, block\}$ in terms of an open net which synchronizes with the original model extended with a configuration interface.

Definition 10 (Canonical configuration partner; allow by default). Let N be an open net and let $C_N : L \rightarrow \{allow, hide, block\}$ be a configuration for N . $Q_a^{C_N} = (P, T, F, m_0, \Omega, L^Q, \ell)$ is the canonical configuration partner with:

- $B = \{x \in L \mid C_N(x) = block\}$ is the set of blocked labels,
- $P = \{p_x^0 \mid x \in B\} \cup \{p_x^\omega \mid x \in B\}$,
- $T = \{t_x \mid x \in B\} \cup \{t_{start}\}$,
- $F = \{(p_x^0, t_x) \mid x \in B\} \cup \{(t_x, p_x^\omega) \mid x \in B\} \cup \{(p_x^\omega, t_{start}) \mid x \in B\}$,

- $m_0 = [(p_x^0)^1 \mid x \in B]$,⁵
- $\Omega = \{ [] \}$,
- $L^Q = \{ \text{block}_x \mid x \in B \} \cup \{ \text{start} \}$,
- $\ell(t_x) = \text{block}_x$ for $x \in B$, $\ell(t_{\text{start}}) = \text{start}$.

The set of labels which need to be blocked to mimic configuration C_N is denoted by B . The canonical configuration partner $Q_a^{C_N}$ has a transition for each of these labels. These transitions may fire in any order after which the transition with label start fires. We observe that in the composition $N_a^{C_I} \oplus Q_a^{C_N}$ first all transitions with a label in $\{ \text{block}_x \mid x \in B \}$ fire in a synchronous manner, and next the transition with label start (in both nets). After this, the net is configured and $Q_a^{C_N}$ plays no role in the composition $N_a^{C_I} \oplus Q_a^{C_N}$ any more.

The following lemma formalizes the relation between the composition $N_a^{C_I} \oplus Q_a^{C_N}$ and feasibility.

Lemma 1. *Let N be an open net and let C_N be a configuration for N . C_N is a feasible configuration iff $N_a^{C_I} \oplus Q_a^{C_N}$ is weakly terminating.*

Proof. (\Rightarrow) Let C_N be a feasible configuration for N and let $N_a^{C_I}$ be as defined in Def. 9. Consider the composition $N_a^{C_I} \oplus Q_a^{C_N}$ after the synchronization via label start has occurred. By construction, (1) $N_a^{C_I} \oplus Q_a^{C_N}$ reached the marking $m = m_0 \oplus m_1 \oplus m_2$ such that m_0 is the initial marking of N , m_1 marks all places p_t^a of transitions $t \in A_N^C \cup H_N^C$, and m_2 is the empty marking of Q^{C_N} . Furthermore, (2) all transitions which bear a synchronization label (i. e., t_{start} and all b_x transitions) and all $t \in B_N^C$ are dead in m and cannot become enabled any more. From $N_a^{C_I}$, construct the net N^* by removing these transitions and their adjacent arcs, as well as the places p_{start} and p_t^a for all $t \in T^V$. The resulting net N^* coincides with β_N^C (modulo renaming). Hence, $N_a^{C_I} \oplus Q_a^{C_N}$ weakly terminates.

(\Leftarrow) Assume $N_a^{C_I} \oplus Q_a^{C_N}$ weakly terminates. From $Q_a^{C_N}$, we can straightforwardly derive a configuration C for N in which all labels are blocked which occur in $N_a^{C_I} \oplus Q_a^{C_N}$. With the same observation as before, we can conclude that β_N^C coincides with the net N^* constructed from $N_a^{C_I}$ after the removal the described nodes. Hence, β_N^C weakly terminates and C is a feasible configuration for N . \square

Lemma 1 states that checking the feasibility of a particular configuration can be reduced to checking for weak termination of the composition. However, the reason for modeling configurations as partners is that we can synthesize partners and test for the existence of feasible configurations.

Theorem 1 (Feasibility coincides with controllability). *Let N be an open net. $N_a^{C_I}$ is controllable iff there exists a feasible configuration C_N of N .*

Proof. (\Rightarrow) If $N_a^{C_I}$ is controllable, then there exists a partner N' of $N_a^{C_I}$ such that $N_a^{C_I} \oplus N'$ is weakly terminating. Consider a marking m of the composition reached by a run σ from the initial marking of $N_a^{C_I} \oplus N'$ to the synchronization

⁵ $[x^k \mid x \in X]$ denotes the multiset where each element of X appears k times. $[]$ denotes the empty multiset.

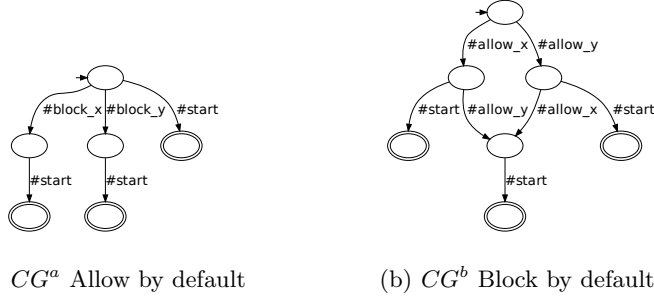


Fig. 4. Two configuration guidelines characterizing all possible configurations.

via label *start*. Using the construction from the proof of Lemma 1, we can derive a net N^* from N_a^{CI} which coincides with a configured net β_N^C for a configuration C_N . As $N_a^{CI} \oplus N'$ is weakly terminating, C_N is feasible.

(\Leftarrow) If C_N is a feasible configuration of N , then by Lemma 1, $N_a^{CI} \oplus Q_a^{C_N}$ weakly terminates and by Def. 8, N_a^{CI} is controllable. \square

As shown in [22], it is possible to synthesize a partner which is *most-permissive*. This partner simulates any other partner and thus characterizes all possible feasible configurations. In previous papers on partner synthesis in the context of service oriented computing, the notion of an *operating guideline* was used to create a finite representation capturing all possible partners [18]. Consequently, we use the term *Configuration Guideline* (CG) to denote the most-permissive partner of a configuration interface. Figure 4(a) shows the configuration guideline CG^a for the configurable model in Fig. 3(a), computed from the configuration interface N_a^{CI} in Fig. 3(b).

A configuration guideline is an automaton with one start state and one or more final states. *Any path in the configuration guideline starting in the initial state and ending in a final state corresponds to a feasible configuration.* The initial state in Fig. 4(a) is denoted by a small arrow and the final states are denoted by double circles. The leftmost path in Fig. 4(a) (i. e., $\langle \text{block}_x, \text{start} \rangle$), corresponds to the configuration which blocks label x . Path $\langle \text{block}_y, \text{start} \rangle$ corresponds to the configuration which blocks label y . The rightmost path (i. e., $\langle \text{start} \rangle$) does not block any label. The three paths capture all three feasible configurations. For example, blocking both labels is not feasible. Figure 4(a) is trivial because there are only two labels and three feasible configurations. However, configuration guidelines can be automatically computed for large and complex configurable process models.

Thus far, we used a configuration interface that allows all configurable activities by default, i. e., blocking is an explicit action of the partner. It is also possible to use a completely different starting point and initially block all activities.

Definition 11 (Configuration interface; block by default). Let $N = (P, T, F, m_0, \Omega, L, \ell)$ be an open net. We define the open net with configuration interface $N_b^{CI} = (P^C, T^C, F^C, m_0^C, \Omega^C, L^C, \ell^C)$ with

- $T^V = \{t \in T \mid \ell(t) \neq \tau\}$,
- $P^C = P \cup \{p_{start}\} \cup \{p_t^a \mid t \in T^V\} \cup \{p_x^b \mid x \in L\}$,
- $T^C = T \cup \{t_{start}\} \cup \{a_x \mid x \in L\}$,
- $F^C = F \cup \{(p_{start}, t_{start})\} \cup \{(t_{start}, p) \mid p \in P \wedge m_0(p) = 1\} \cup \{(t, p_t^a) \mid t \in T^V\} \cup \{(p_t^a, t) \mid t \in T^V\} \cup \{(a_x, p_{start}) \mid x \in L\} \cup \{(p_{start}, a_x) \mid x \in L\} \cup \{(a_{\ell(t)}, p_t^a) \mid t \in T^V\} \cup \{(p_x^b, a_x) \mid x \in L\}$,
- $m_0^C = [p^1 \mid p \in \{p_{start}\} \cup \{p_x^b \mid x \in L\}]$,
- $\Omega^C = \{m \oplus [(p_x^b)^1 \mid x \in X] \oplus [(p_t^a)^1 \mid t \in T \wedge \ell(t) \notin X] \mid m \in \Omega \wedge X \subseteq L\}$,
- $L^C = \{start\} \cup \{allow_x \mid x \in L\}$
- $\ell^C(t_{start}) = start$, $\ell^C(a_x) = allow_x$ for $x \in L$, and $\ell^C(t) = \tau$ for $t \in T$.

N_b^{CI} in Fig. 3(c) shows the configuration interface where all activities are blocked by default. The idea is analogous to the construction of N_a^{CI} . Transitions a_x and a_y are added to model the allowing of labels x and y , and places p_t^a ($t \in T^V$) are added to connect these transitions to the original ones. However these places are empty, and thus all original transitions are initially blocked; that is, they cannot fire. An original transition (e.g., x) can only be enabled after its allowing transition (i.e., a_x) fires. Places p_x^b and p_y^b have been added to inhibit the repeated execution of respectively a_x and a_y . Without these places, the inner net (i.e., the net without synchronization labels) would be unbounded, and controllability would be undecidable [20]. Similar to the “allow by default” case, we define a canonical configuration partner.

Definition 12 (Canonical configuration partner; block by default). *Let N be an open net and let $C_N : L \rightarrow \{allow, hide, block\}$ be a configuration for N . $Q_b^{CN} = (P, T, F, m_0, \Omega, L^Q, \ell)$ is the canonical configuration partner with:*

- $A = \{x \in L \mid C_N(x) \neq block\}$ is the set of nonblocked labels,
- $P = \{p_x^0 \mid x \in A\} \cup \{p_x^\omega \mid x \in A\}$,
- $T = \{t_x \mid x \in A\} \cup \{t_{start}\}$,
- $F = \{(p_x^0, t_x) \mid x \in A\} \cup \{(t_x, p_x^\omega) \mid x \in A\} \cup \{(p_x^\omega, t_{start}) \mid x \in A\}$,
- $m_0 = [(p_x^0)^1 \mid x \in A]$,
- $\Omega = \{ [] \}$,
- $L^Q = \{allow_x \mid x \in A\} \cup \{start\}$,
- $\ell(t_x) = allow_x$ for $x \in A$, $\ell(t_{start}) = start$.

The structure of the canonical configuration partner Q_b^{CN} is identical to that of Q_a^{CN} . Only the labels are different; that is, $A = L \setminus B$ are the labels that need to be unblocked. Moreover, we obtain the same results linking feasibility to controllability.

Lemma 2. *Let N be an open net and let C_N be a configuration for N . C_N is a feasible configuration iff $N_b^{CI} \oplus Q_b^{CN}$ is weakly terminating.*

Proof. Analogue to the proof of Lemma 1. □

Theorem 2 (Feasibility coincides with controllability). *Let N be an open net. N_b^{CI} is controllable iff there exists a feasible configuration C_N of N*

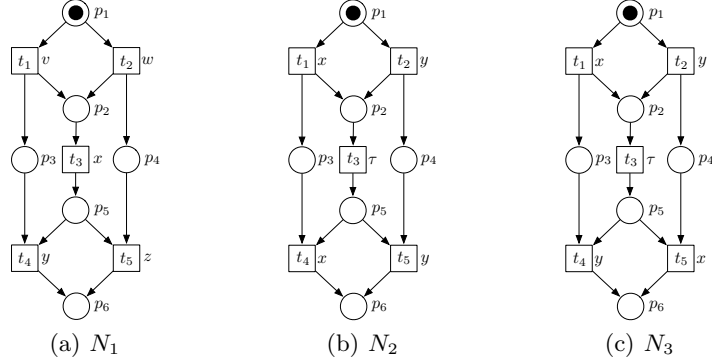


Fig. 5. Three open nets ($\Omega = \{[p_6]\}$).

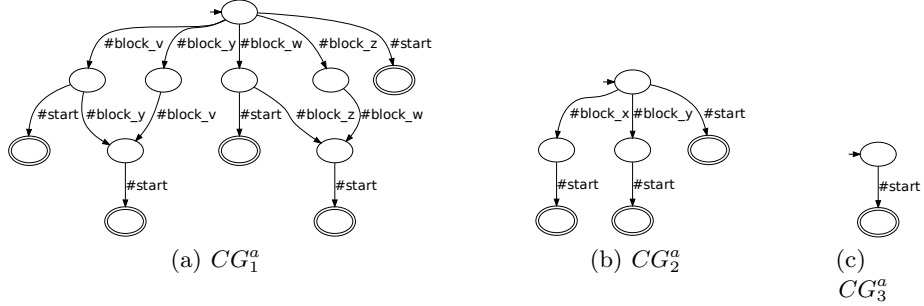


Fig. 6. The configuration guidelines (allow by default) for N_1 (a), N_2 (b) and N_3 (c).

Proof. Analogue to the proof of Theorem 1. \square

Figure 4(b) shows the configuration guideline CG^b for the configurable model in Fig. 3(a), computed from the configuration interface N_b^{CI} in Fig. 3(c). Again, any path in CG^b starting in the initial state and ending in a final state correspond to a feasible configuration. The leftmost path (i. e., $\langle \text{allow}_x, \text{start} \rangle$) corresponds to the configuration which unblocks label x . Paths $\langle \text{allow}_x, \text{allow}_y, \text{start} \rangle$ and $\langle \text{allow}_y, \text{allow}_x, \text{start} \rangle$ correspond to the configuration where both x and y are allowed. Finally, the rightmost path (i. e., $\langle \text{allow}_y, \text{start} \rangle$) allows y only. Clearly, the two configuration guidelines in Fig. 4 point to the same set of feasible configurations as they refer to the same original model.

Let us now consider a more elaborated example to see how configuration guidelines can be used to rule out unfeasible configurations. Figure 5 shows three open nets. The structures are identical, only the labels are different. For example, blocking x in N_2 corresponds to removing both t_1 and t_4 , because both transitions bear the same label.

For these three nets, we can construct the configuration interfaces using Def. 9 or Def. 11, and then synthesize the configuration guidelines. Figure 6 shows the three configuration guidelines using Def. 9 (allow by default).

Figure 6(a) reveals all feasible configurations for N_1 in Fig. 5(a). From the initial state in the configuration guideline CG_1^a , we can immediately reach a final state by following the rightmost path $\langle \text{start} \rangle$. This indicates that all configurations which block nothing (i. e., only allow or hide activities) are feasible. It is possible to just block v (cf. path $\langle \text{block}_v, \text{start} \rangle$) or block both v and y (cf. paths $\langle \text{block}_v, \text{block}_y, \text{start} \rangle$ and $\langle \text{block}_y, \text{block}_v, \text{start} \rangle$). However, as Fig. 6(a) shows, it is not allowed to block y only, otherwise a token would deadlock in p_3 . For the same reasons, one can block w only or w and z , but not z only. Moreover, it is not possible to combine the blocking of w and/or z on the one hand and v and/or y on the other hand, otherwise no final marking can be reached. Also x can never be blocked, otherwise both v and w would also need to be blocked (to avoid a token to deadlock in p_2) which is not possible. There are $3^5 = 243$ configurations for N_1 . If we abstract from hiding as this does not influence feasibility, there remain $2^5 = 32$ possible configurations. Of these only 5 are feasible configurations which correspond to the final states in Fig. 6(a). This illustrates that the configuration guideline can indeed represent all feasible configurations in an intuitive manner.

Figure 6(b) shows the three feasible configurations for N_2 in Fig. 5(b). Again all final states correspond to feasible configurations. As the configuration guideline shows one can block x or y but not both. It is easy to see that one can indeed block the two leftmost transitions (labeled x) or the two rightmost transitions (labeled y), but not both.

The configuration guideline in Fig. 6(c) shows that nothing can be blocked for N_3 (Fig. 5(c)). Blocking x or y will yield an unfeasible configuration as a token will get stuck in p_4 (when blocking x) or p_3 (when blocking y). If both are blocked, none of the transitions can fire and thus no final marking can be reached.

5 Tool Support

To prove the feasibility of our approach, we applied it to the configuration of C-YAWL models [13]. The YAWL language can be seen as an extension of Petri nets which provides “syntactic sugaring” (shorthand notations for sequences and XOR-splits/joins) and advanced constructs such as cancelation sets, multiple instance tasks and OR-joins [14]. YAWL is based on the well-know workflow patterns [4]. The YAWL system supporting this language is one of the most widely used open source workflow systems [14]. Here we do not use YAWL’s cancelation sets, multiple instance tasks and OR-joins and restrict ourselves to the basic control-flow patterns supported by most systems. This allows us to easily map a YAWL model onto an open net.

A C-YAWL model is a YAWL model where some tasks are annotated as *configurable*. Configuration is achieved by restricting the routing behavior of configurable tasks via the notion of *ports*. A configurable task’s joining behavior is identified by one or more *inflow* ports, whereas its splitting behavior is identified by one or more *outflow* ports. The number of ports for a configurable task

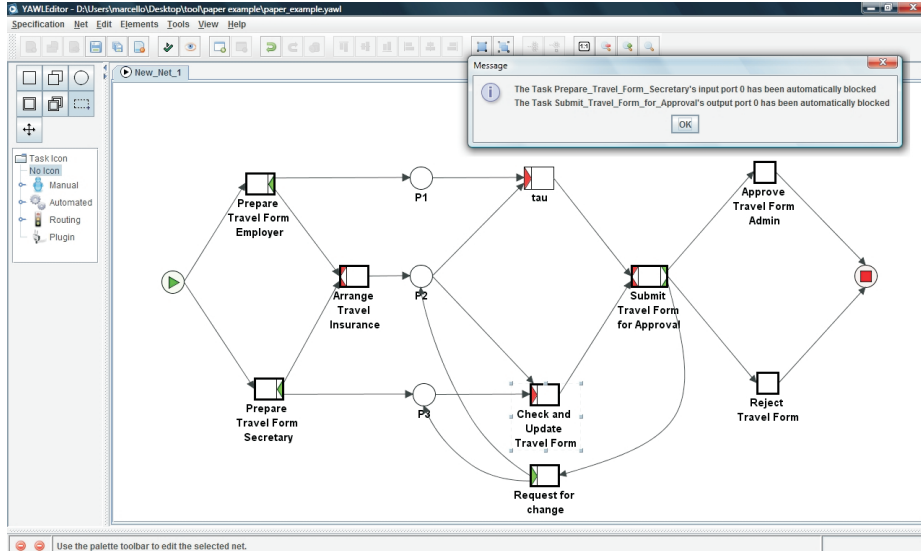


Fig. 7. The C-YAWL model for travel request approval.

depends on the task's routing behavior. For example, an AND-split/join and an OR-join are each identified by a single port, whereas an XOR-split/join is identified by one port for each outgoing/incoming flow. An OR-split is identified by a port for each combination of outgoing flows. To restrict a configurable task's routing behavior, inflow ports can be hidden (thus the corresponding task will be skipped) or blocked (no control will be passed to the corresponding task via that port), whereas outflow ports can only be blocked (the outgoing paths from that task via that port are disabled). For instance, Fig. 7 shows the C-YAWL model for the travel request approval in the YAWL Editor, where configurable tasks are marked with a ticker border.

The YAWL Editor offers a visual interface to conveniently configure C-YAWL models and obtain configured models. Given a configuration, the tool can show a preview of the resulting configured net by greying out all model fragments which have been blocked, and commit the configuration by removing these fragments altogether. To assist end users in ruling out all unfeasible configurations in an interactive manner, we developed an open-source plugin for the YAWL Editor named *C-YAWL Correctness Checker*.⁶ Given a C-YAWL model in memory, the plugin first maps this model into an open net. More precisely, it maps each condition to a place, each configurable task's port to a labeled transition, and each non-configurable task to a silent transition. Also, for each task it adds an extra place to connect the transition(s) derived from its inflow port(s) with the transition(s) derived from its outflow port(s). By using silent transitions we prevent all non-configurable tasks from being later configured via a configuration

⁶ The C-YAWL Correctness Checker (and the C-YAWL system) can be downloaded from www.yawlfoundation.org.

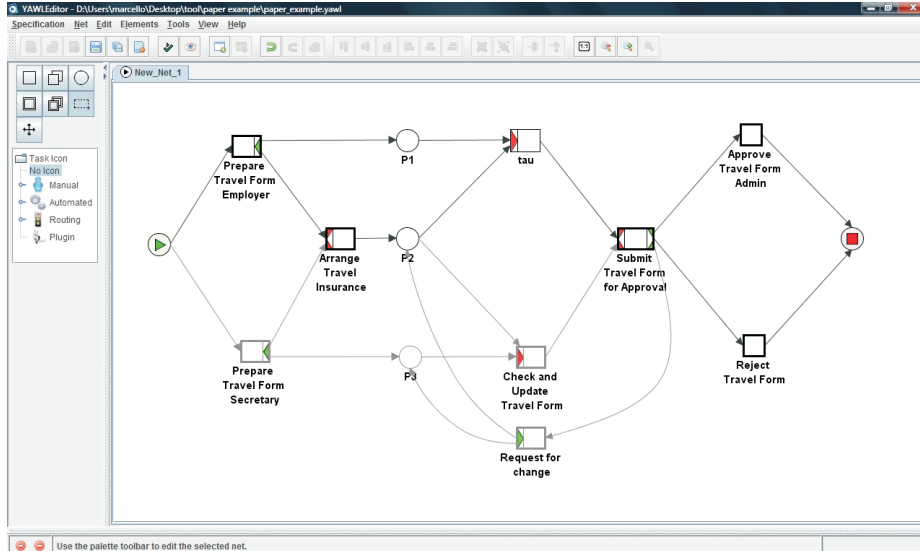


Fig. 8. The preview of a configured net for the example in Fig. 7.

interface. Next, the plugin uses the generated open net to create the corresponding configuration interface (allow by default), and passes the latter to the tool *Wendy* [19] to produce the configuration guideline (allow by default). *Wendy* is an open source tool⁷ which implements the algorithms for partner synthesis [22] and to calculate operating guidelines [18]. A recent case study [19], shows that *Wendy* is able to analyze industrial models with up to 5 million states and to synthesize partners of about the same size. *Wendy* itself offers no graphical user interface, but is controlled by input/output streams. In our setting, *Wendy*'s output is piped back into the Correctness Checker, where it can be parsed.

At each configuration step, the plugin scans the set of outgoing edges of the current state in the configuration guideline, and prevents users from blocking those ports not included in this set. This is done by disabling the block button for those ports. As users block a valid port, the Correctness Checker traverses the configuration guideline through the corresponding edge and updates the current state. If this is not a *consistent* state; that is, a state with an outgoing edge labeled “start”, further ports need to be blocked, because the current configuration is unfeasible. In this case *YAWL* provides an “*auto complete*” option. This is achieved by traversing the shortest path from the current state to a consistent state and automatically blocking all ports in that path. After this, the plugin notifies the user with the list of ports being blocked and updates the current state. For example, Fig. 7 shows that after blocking the input port of task *Check and Update Travel Form*, the plugin notifies the user that the input port of task *Prepare Travel Form for Approval (Secretary)* and the output port of task *Submit Travel Form for Approval* to task *Request for Change* have also been blocked.

⁷ Available for download at <http://service-technology.org/wendy>.

Figure 8 shows the preview of the resulting configured net. From this we can observe that condition p_3 and task *Request for Change* will also be removed from the net as a result of applying the earlier configuration.

Similarly, the plugin maintains a consistent state in case users decide to allow a previously blocked port. In this case it traverses the shortest backward path to a consistent state and allows all ports in that path. By traversing the shortest path we ensure that the number of ports being automatically blocked or allowed is minimal.

The C-YAWL example of Fig. 7 comprises ten inflow ports and nine outflow ports. In total more than 30 million configurations are potentially possible. If we abstract from hiding we obtain 524,288 possible configurations, of which only 1,593 are feasible according to the configuration guideline. Wendy took an average of 336 seconds to generate this configuration guideline which consumes 3.37 MB of disk space. Nonetheless, the shortest path computation is a simple depth-first search which is linear on the number of nodes in the configuration guideline. Thus, once the configuration guideline has been generated, the plugin's response time at each user interaction is instantaneous.

6 Conclusion

Configurable process models are a means to compactly represent families of process models. However, the verification of such models is difficult as the number of possible configurations grows exponentially in the number of configurable elements. Due to concurrency and branching structures, configuration decisions may interfere with each other and thus introduce deadlocks, livelocks and other anomalies. The verification of configurable process models is challenging and only few researchers have worked on this. Moreover, existing results impose restrictions on the structure of the configurable process model and fail to provide insight into the complex dependencies between configuration decisions.

This paper uses an innovative approach where configuration is seen as an external service. This service acts as a partner which can block or allow particular activities. Using partner synthesis we compute the configuration guideline – a compact representation characterizing all external services that yield the desired behavior, which correspond to all feasible configurations. The approach is highly generic and imposes no constraints on the configurable process model. Moreover, all computations are done at design time and not at configuration time. As a result, once the configuration guideline has been generated, the response time is instantaneous thus stimulating the practical (re-)use of configurable process models. The approach is supported by a combination of the YAWL system and Wendy. As a result, C-YAWL models can be configured while correctness is ensured by a checker integrated in the YAWL Editor.

Several interesting extensions are possible. First, the partner synthesis could be further refined using *behavioral constraints* [17]. With such constraints, specific partners can be ruled out. This could be used to encode knowledge about a process' application domain [16] in the configuration interface. For example,

domain knowledge may state that two activities cannot be blocked or allowed at the same time. Second, one could consider configuration at run-time, i. e., while instances are running configurations can be set or modified. This can be easily embedded in the current approach. Finally, one could devise even more compact representations of configuration guidelines. For example, the diamond structure in Fig. 4(b) suggests that this configuration guideline could be represented more efficiently. An idea would be to convert the automaton into a Petri net using the theory of regions [5].

References

1. W.M.P. van der Aalst. The Application of Petri Nets to Workflow Management. *The Journal of Circuits, Systems and Computers*, 8(1):21–66, 1998.
2. W.M.P. van der Aalst and T. Basten. Inheritance of Workflows: An Approach to Tackling Problems Related to Change. *Theoretical Computer Science*, 270(1-2):125–203, 2002.
3. W.M.P. van der Aalst, M. Dumas, F. Gottschalk, A.H.M. ter Hofstede, M. La Rosa, and J. Mendling. Preserving Correctness During Business Process Model Configuration. *Formal Aspects of Computing*, 2010.
4. W.M.P. van der Aalst, A.H.M. ter Hofstede, B. Kiepuszewski, and A.P. Barros. Workflow Patterns. *Distributed and Parallel Databases*, 14(1):5–51, 2003.
5. E. Badouel and P. Darondeau. Theory of regions. In *Advanced Course on Petri Nets*, LNCS 1491, pages 529–586. Springer, 1996.
6. T. Basten and W.M.P. van der Aalst. Inheritance of Behavior. *Journal of Logic and Algebraic Programming*, 47(2):47–145, 2001.
7. J. Becker, P. Delfmann, and R. Knackstedt. Adaptive Reference Modeling: Integrating Configurative and Generic Adaptation Techniques for Information Models. In *Reference Modeling: Efficient Information Systems Design Through Reuse of Information Models*, pages 27–58. Physica-Verlag, Springer, 2007.
8. T. Curran and G. Keller. *SAP R/3 Business Blueprint: Understanding the Business Process Reference Model*. Upper Saddle River, 1997.
9. K. Czarnecki and M. Antkiewicz. Mapping Features to Models: A Template Approach Based on Superimposed Variants. In *GPCE 2005*, pages 422–437. Springer, 2005.
10. J. Desel and J. Esparza. *Free Choice Petri Nets*, volume 40 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, Cambridge, UK, 1995.
11. P. Fettke and P. Loos. Classification of Reference Models - A Methodology and its Application. *Information Systems and e-Business Management*, 1(1):35–53, 2003.
12. F. Gottschalk, W.M.P. van der Aalst, and H.M. Jansen-Vullers. Configurable Process Models: A Foundational Approach. In *Reference Modeling: Efficient Information Systems Design Through Reuse of Information Models*, pages 59–78. Physica-Verlag, Springer, 2007.
13. F. Gottschalk, W.M.P. van der Aalst, M.H. Jansen-Vullers, and M. La Rosa. Configurable Workflow Models. *Int. J. Cooperative Inf. Syst.*, 17(2):177–221, 2008.
14. A.H.M. ter Hofstede, W.M.P. van der Aalst, M. Adams, and N. Russell. *Modern Business Process Automation: YAWL and its Support Environment*. Springer, 2010.
15. M. La Rosa, M. Dumas, A.H.M. ter Hofstede, J. Mendling, and F. Gottschalk. Beyond Control-Flow: Extending Business Process Configuration to Roles and Objects. In *ER 2008*, volume 5231 of *LNCS*, pages 199–215. Springer, 2008.

16. M. La Rosa, J. Lux, S. Seidel, M. Dumas, and A.H.M. ter Hofstede. Questionnaire-driven Configuration of Reference Process Models. In *CAiSE'07*, volume 4495 of *LNCS*, pages 424–438. Springer, 2007.
17. N. Lohmann, P. Massuthe, and K. Wolf. Behavioral Constraints for Services. In *BPM 2007*, volume 4546 of *LNCS*, pages 271–287. Springer, 2007.
18. N. Lohmann, P. Massuthe, and K. Wolf. Operating Guidelines for Finite-State Services. In *ICATPN 2007*, volume 4546 of *LNCS*, pages 321–341. Springer, 2007.
19. N. Lohmann and D. Weinberg. Wendy: A tool to synthesize partners for services. In *PETRI NETS 2010*, LNCS. Springer, 2010. Available for download at <http://service-technology.org/wendy>.
20. P. Massuthe, A. Serebrenik, N. Sidorova, and K. Wolf. Can I find a Partner? Undecidability of Partner Existence for Open Nets. *Information Processing Letters*, 108(6):374–378, 2008.
21. M. Rosemann and W.M.P. van der Aalst. A Configurable Reference Modelling Language. *Information Systems*, 32(1):1–23, 2007.
22. K. Wolf. Does my service have partners? *LNCS T. Petri Nets and Other Models of Concurrency*, 5460(2):152–171, 2009.
23. YAWL Foundation. Home Page. <http://www.yawlfoundation.org>. Accessed: March 2010.