Electronic Communications of the EASST
Volume 076 (2019)

Automated Verification of Critical Systems 2018 (AVoCS 2018)

Using an SMT engine to generate Symbolic Automata

Xudong Qin, Simon Bliudze, Eric Madelaine, and Min Zhang

17 pages

Guest Editors: David Pichardie, Mihaela Sighireanu
ECEASST Home Page: http://www.easst.org/eceasst/ ISSN 1863-2122

Using an SMT engine to generate Symbolic Automata

Xudong Qin[1]*, Simon Bliudze[2], Eric Madelaine[3], and Min Zhang[4]

[1] steven_qxd@126.com
[4] mzhang@sei.ecnu.edu.cn
Shanghai Key Laboratory of Trustworthy Computing, ECNU, China

[2] simon.bliudze@inria.fr
INRIA Lille – Nord Europe, 40 avenue Halley, 59650 Villeneuve d'Ascq, France

[3] eric.madelaine@inria.fr
Université Côte d'Azur, Inria, CNRS, I3S, 06902 Sophia Antipolis, France

Abstract: Open pNets are used to model the behaviour of open systems, both synchronous or asynchronous, expressed in various calculi or languages. They are endowed with a symbolic operational semantics in terms of so-called "Open Automata". This allows us to check properties of such systems in a compositional manner. We implement an algorithm computing these semantics, building predicates expressing the synchronization conditions between the events of the pNet sub-systems. Checking such predicates requires symbolic reasoning over first order logics, but also over application-specific data. We use the Z3 SMT engine to check satisfiability of the predicates, and prune the open automaton of its unsatisfiable transitions. As an industrial oriented use-case, we use so-called "architectures" for BIP systems, that have been used in the framework of an ESA project and to specify the control software of a nanosatellite at the EPFL Space Engineering Center. We use pNets to encode a BIP architecture extended with explicit data, and compute its open automaton semantics. This automaton may be used to prove behavioural properties; we give 2 examples, a safety and a liveness property.

Keywords: Networks of Synchronised Automata, Symbolic Behavioral Semantics, Compositional Analysis of Software Components, SMT

1 Introduction

In the nineties, several works extended the basic behavioural models based on labelled transition systems to address value-passing or parameterized systems, using various symbolic encodings of the transitions [1, 2, 3, 4]. In [4], H.M. Lin addressed value-passing calculi, for which he developed a symbolic behavioural semantics, and proved algebraic properties. Separately J. Rathke [5] defined another symbolic semantics for a parameterized broadcast calculus, together with strong and weak bisimulation equivalences, and developed a symbolic model-checker based on a tableau method for these processes. Thirty years later, no practical verification approach and

* This work was partially funded by the Associated Team FM4CPS between INRIA and ECNU, Shanghai

no verification platform are using this kind of approaches to provide proof methods for value-passing processes or open process expressions.

Parameterized Networks of Synchronized Automata (pNets) were proposed to give a behavioural specification formalism for distributed systems, synchronous, asynchronous, or heterogeneous. They are used in VerCors [6], a platform for designing and verifying distributed systems, as the intermediate language for various high-level languages. The high-level languages in VerCors formalize each component of the distributed system and their composition. pNets provides the core low-level semantic formalism for VerCors, and is made of a hierarchical composition of (value-passing) automata, called parameterized labelled transition systems (pLTS), where each hierarchical level defines the possible synchronization of the lower levels. Traditionally, pNets have been used to formalize fully defined systems or softwares. But we want also to define and reason about incompletely defined systems, like program skeletons, operators, or open expressions of process calculi. The open pNet model addresses this problem, using "holes" as process parameters, representing *unspecified* subsystems. The pNet model was developed in a series of papers [7, 8] in which many examples have been introduced showing its ability to encode the operators from some other algebras or program skeletons. The operational semantics of an (open) pNet is defined as an Open Automaton in which Open Transitions contain logical predicates expressing the relations between the behaviour of the holes, and the global behaviour of the system. In the previous publication, only a sketch of a procedure allowing to compute these semantics was presented, together with a proof of finiteness of the open automaton, under reasonable hypotheses on the pNet structure.

Implementing these semantics raised several challenges, in order:

- to get a tool that could be applied to pNets representing various languages, in particular various actions algebras, with their specific decision theories,
- to separate clearly the algorithm generating the transitions of the open automaton from combination of all possible (symbolic) behaviours, from the symbolic reasoning part, specifically here using an SMT engine to check the satisfiability of the predicates generated by our algorithm,
- to build a prototype and validate the approach on our basic case-studies, and understand the efficiency of the interaction with the SMT solver.

In the long term, we want to be able to check the equivalence between open systems encoded as pNets. The equivalence between pNets is "FH-bisimulation" [8], a dedicated version of symbolic bisimulation taking the predicate of the open transitions into account when matching such open transitions. We foresee that the interplay with the SMT solver that we use here for satisfiability of open transitions will be similar with what we need when proving (symbolic) equivalence between open transitions.

Contribution In the article we show how:

- We define the open automaton generation algorithm. We implemented a full working prototype, within the VerCors platform. In the process, we improved the semantics rules from [8], and add features in the algorithm to deal with the full model, including management of variables and assignments.
- We implement the interaction between our algorithm and the Z3 SMT solver, for checking

satisfiability of the transitions generated by the algorithm.
- We show the interest of this approach on an industry-inspired case-study, namely one architectural pattern extracted (and extended) from the BIP specification of a nanosatellite on-board software.

Related work. Very few attempts were made to develop symbolic bisimulation approaches for the value-passing process algebras and languages—our long-term goals—especially, there is no algorithmic treatment of the symbolic systems developed by interacting with automatic theorem provers. The closest work is the one already mentioned from J. Rathke [5], who developed the symbolic bisimulation for a calculus of broadcasting system (CBS). CBS is similar with classic process calculi such as CCS and CSP, but communicating by broadcasting values, transmitting values without blocking. That makes the definition of the symbolic semantic and bisimulation equivalence different from the classic works.

For other applications, such as the analysis of programming languages, there exist dedicated platforms using external automatic theorem provers (ATP) or automatic tactics from interactive theorem provers (ITP), to perform symbolic reasoning, and for example to discharge some subgoals in the proofs. Tools like Rodin [9, 10] have already integrated several provers, like Z3, as modules for proving the proof obligations generated from a user model. The prover we use, which also happens to be Z3, is developed by Microsoft Research based on the satisfiability modulo theories framework (SMT), is mainly applied in extended static checking, test case generation, and predicate abstraction. In a similar way, there are several ATPs/ITPs we could consider to use for result pruning and bisimulation checking in our algorithm, as an alternative to Z3, such as CVC4 [11], Coq [12] or others.

BIP (Behaviour-Interaction-Priority) [13] is a framework for the component-based design of concurrent software and systems. In particular, the BIP tool-set comprises compilers for generating C/C++ code, executable by linking with one of the dedicated engines, which implement the BIP operational semantics [14]. This approach ensures that any property, shown to hold on a given BIP model, will also hold by construction on the generated code. BIP Architectures [15] formalise design patterns, which enforce global properties characterising the coordination among the components of the system. They provide a compositional approach, ensuring correctness by construction during the design of BIP models. In [15], it was shown that application of architectures is compositional w.r.t. safety properties, i.e. when several architectures are applied, each enforcing a safety property, the resulting system satisfies their conjunction.

But the interaction feature in architectures does not handle data-sensitive interaction constraints. Using an encoding of architectures, extended with data-dependant interactions, into open pNets was an interesting alternative to a direct extension to the architecture semantics.

Structure. In section 2 we give a description and a formal definition of the pNet model, as found in previous publications. Then in section 3 we present our use-case, based on a BIP architecture from the nano-satellite case-study. Section 4 recalls briefly the operational semantics of pNet. Section 5 explains in details the algorithm used to compute this semantics, including the interaction with Z3, and shows the full result of the semantic computation on the running example. Finally we conclude and discuss perspectives in Section 6.

2 Background: pNets definition

This section introduces pNets and the notations we will use in this paper. Then it gives the formal definition of pNet structures, together with an operational semantics for open pNets.

pNets are tree-like structures, where the leaves are either *parameterized labelled transition systems (pLTSs)*, expressing the behaviour of basic processes, or *holes*, used as placeholders for unknown processes, of which we only specify their set of possible actions, named *sort*. Nodes of the tree (pNet nodes) are synchronizing artifacts, using a set of *synchronization vectors* that express the possible synchronization between the parameterized actions of a subset of the sub-trees.

Notations. We extensively use indexed structures over some countable indexed sets, which are equivalent to mappings over the countable set. $a_i^{i\in I}$ denotes a family of elements a_i indexed over the set I. When this is not ambiguous, we shall use notations for sets, and typically write "indexed set over I" when formally we should speak of multisets, and write $x \in a_i^{i\in I}$ to mean $\exists i \in I. x = a_i$. An empty family is denoted $\emptyset$. We denote classically $\overline{a}$ a family when the indexing set is irrelevant. $\uplus$ is the disjoint union on indexed sets.

Term algebra. Our models rely on a notion of parameterized actions that are symbolic expressions using data types and variables. As we want to encode the low-level behaviour of possibly very different programming languages, we do not want to impose one specific algebra for denoting actions, nor any specific communication mechanism. So we leave unspecified the constructors of the algebra that will allow building expressions and actions. Moreover, we use a generic *action interaction* mechanism, based on unification between two or more action expressions. This will be used in the semantics of synchronization vectors to express various kinds of communication or synchronization mechanisms.

Formally, we assume the existence of a term algebra $\mathscr{T}_{\Sigma,\mathscr{V}}$, where Σ is the signature of the data and action constructors, and $\mathscr{V}$ a set of variables. Within $\mathscr{T}_{\Sigma,\mathscr{V}}$, we distinguish a set of data expressions $\mathscr{E}_{\mathscr{V}}$, including a set of Boolean expressions $\mathscr{B}_{\mathscr{V}}$ ($\mathscr{B}_{\mathscr{V}} \subseteq \mathscr{E}_{\mathscr{V}}$). On top of $\mathscr{E}_{\mathscr{V}}$ we build the action algebra $\mathscr{A}_{\mathscr{V}}$, with $\mathscr{A}_{\mathscr{V}} \subseteq \mathscr{T}_{\mathscr{V}}, \mathscr{E}_{\mathscr{V}} \cap \mathscr{A}_{\mathscr{V}} = \emptyset$; naturally action terms will use data expressions as sub-terms. The function $vars(t)$ identifies the set of variables in a term $t \in \mathscr{T}$.

pNets can encode naturally the notion of input actions as found, e.g. in value-passing CCS [16] or of usual point-to-point message passing calculi, but it also allows for more general mechanisms, like gate negotiation in Lotos [17], or broadcast communications.

Algebra presentations. In practice, the parameterization of the pNet model by some specific action algebra is realized by the definition of a many-sorted "algebra presentation". It will be used to check the well-formedness of a pNet system, and to define the translation of the pNet semantics into the SMT engine input language ([18]).

Definition 1 An *algebra presentation* is a triple $\mathscr{P} = \langle Sorts, Constrs, Ops \rangle$, where

- *Sorts* is a set of data *sorts*
- *Constrs* is a set of *constructor operators*: for each $Con \in Constrs$, $arity(Con) = n \in \mathbb{N}$ is its arity and '$Con : (sel_1, sort_1), \ldots, (sel_n, sort_n) \rightarrow sort$' is its signature with the associated

Table 1: Algebra Presentation: predefined Sorts and Operators

Sort	Constructors	Auxiliary Operators
Bool	`true`, `false`	$\wedge$, $\vee$, $\neg$, $\Longrightarrow$, $=$, $\neq$
Action	`Synchro`, `FUN`	
Int	$0, \{i, -i\}_{i \in \mathtt{Nat}}$	$-$ (`unary`), $+$, $-$ (`binary`), $\times$, $\div$ etc.
Extension for the BIP use-case of Fig. 2		
Action	FUN_Action_Bool, fail, resume, timeout, reset, start, tick, ask	

selectors. For each argument, the pair $(sel_i, sort_i)$ defines an auxiliary operator of name sel_i with signature $sel_i : sort \to sort_i$.

- Ops is a set of *auxiliary operators*, with their arity and signature, of the form: $Op : sort_1, \ldots, sort_n \to sort$
- $Constrs(sortname)$ and $Sels(sortname)$ are, respectively, the sets of constructors and selectors of the sort *sortname*

Constructors of arity 0 are called *constants*, and denoted $Consts(\mathscr{P})$.

Sorts `Bool` and `Int` are predefined with standard operators. Sort `Action` also, with a constructor `Synchro` denoting a synchronized action, i.e. an "internal" action that cannot be further synchronized with the environment. It also comes with an overloaded `FUN` constructor, used to build actions with arguments, that will be instantiated to the required sorts for a given pNet.

The definition of an Algebra Presentation, and a set of variables $\mathscr{V}$ fixes the Term algebra elements $\mathscr{T}_{\Sigma,\mathscr{V}}, \mathscr{B}_{\mathscr{V}}, \mathscr{A}_{\mathscr{V}}$.

2.1 The (open) pNets Core Model

A pLTS is a labelled transition system with variables, which can be manipulated, defined, or accessed inside states, actions, guards, and assignments. Each state has its set of variables called *state variables*, which can only be modified by the assignment in transitions targeting this state. A global state variable of a pLTS is a state variable defined in all states. Note that we make no assumptions on finiteness of the set of states, nor on finite branching of the transition relation.

We first define the set of actions a pLTS can use. Let a range over action labels, op are operators, and x_i range over variable names. Action terms are:

$$\begin{array}{rcll} \alpha \in \mathscr{A} & ::= & a(p_1, \ldots, p_n) & \text{action terms} \\ p_i & ::= & Expr & \text{parameters} \\ Expr & ::= & Value \mid x \mid op(Expr_1, \ldots, Expr_n) & \text{expressions} \end{array}$$

Definition 2 (pLTS) Given a term algebra $\mathscr{T}_{\Sigma,\mathscr{V}}$, a pLTS is a tuple $pLTS \triangleq \langle\!\langle S, s_0, \to \rangle\!\rangle$ where:

- S is a set of *states*, with $s_0 \in S$ the *initial state*.
- $\to \subseteq S \times L \times S$ is the *transition relation*, with L the set of labels of the form $\langle \alpha,\ e_b,\ (x_j := e_j)^{j \in J} \rangle$, where $\alpha \in \mathscr{A}_{\mathscr{V}}$ is a parameterized action, $e_b \in \mathscr{B}_{\mathscr{V}}$ is a guard, and expressions $\mathscr{E}_{\mathscr{V}} \cup \mathscr{A}_P$ are assigned to x_j. If $s \xrightarrow{\langle \alpha,\ e_b,\ (x_j := e_j)^{j \in J} \rangle} s'$ then $vars(e_b) \subseteq vars(s) \cup vars(\alpha)$, and $\forall j \in J.\, vars(e_j) \subseteq vars(s) \wedge x_j \in vars(s')$.

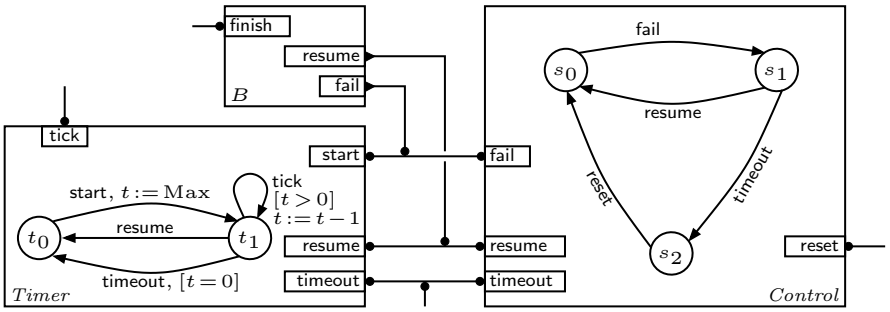

Figure 1: The BIP specification of the Failure Monitor architecture

Now, we define pNet nodes as constructors for hierarchical structures. A pNet node has a set of sub-pNets that can be either pNets or pLTSs, and a set of *holes*, playing the role of process parameters (i.e. unknown in the environment).

A composite pNet consists of a set of sub-pNets, each exposing a set of actions. The relation between actions of a pNet and those of its sub-pNets are given by *synchronization vectors*, which synchronize one or several internal actions, and expose a single resulting global action.

Definition 3 (pNets) A pNet is a hierarchical structure where leaves are pLTSs and holes: $pNet \triangleq pLTS \mid \langle\!\langle pNet_i^{i\in I}, J, SV_k^{k\in K} \rangle\!\rangle$, with I, J, K potentialy infinite, where

- $pNet_i^{i\in I}$ is the family of sub-pNets;
- J is a set of indexes, called *holes*. I and J are *disjoint*: $I \cap J = \emptyset$, $I \cup J \neq \emptyset$
- $SV_k^{k\in K}$ is a set of synchronisation vectors ($K \in \mathscr{I}_{\mathscr{V}}$). $\forall k \in K, SV_k = \alpha_l^{l\in I_k \uplus J_k} \rightarrow \alpha'_k \mid g_k$, where $\alpha'_k \in \mathscr{A}_{\mathscr{V}}$, $I_k \subseteq I$, $J_k \subseteq J$, and $vars(\alpha'_k) \subseteq \bigcup_{l\in I_k \uplus J_k} vars(\alpha_l)$. The global action of a vector SV_k is α'_k. The Boolean expression g_k, such that $vars(g_k) \subseteq \bigcup_{l\in I_k \uplus J_k} vars(\alpha_l)$, is a guard associated to the vector.

In Fig. 2, we show examples of these constructs, with two pLTSs, one hole and one pNet node encoding our running example.

3 Running example

As a running example we use the Failure Monitor architecture from the CubETH nanosatellite on-board software case-study [19] realised using BIP. The architecture-based design process in BIP takes as input a set of components providing basic functionality of the system and a set of temporal properties that must be enforced in the final system. For each property, a corresponding architecture is identified and applied to the model, thereby potentially introducing additional coordinator components and modifying the connectors that define synchronisation patterns among ports of components.

Figure 1 shows a refined version of the Failure Monitor architecture used in [19]. Contrary to standard BIP models, architectures comprise one or several *operand* components, whereof only the set of *ports* is given. Here, the operand component is B and its interface consists of the ports finish, resume and fail. The two *coordinator* components—C and T—are standard BIP components insofar as they also have their *behaviour* specified by finite automata extended with

local data variables. Transitions of these automata are labelled with the ports of the corresponding components, Boolean guards and update functions on local variables. For instance the loop transition $t_2 \xrightarrow{\mathsf{tick},[t<z.u],t:=t+1} t_2$ in the T component is labeled by the port tick, it can be fired only when the current value of the local variable t is greater than 0. Upon firing, this transition decrements the value of t by 1. When omitted, the default guard (resp. update function) is the constant predicate *true* (resp. the *skip* operator). The constant Max, in $t_1 \xrightarrow{\mathsf{s},t:=0,z:=\top} t_2$, is a parameter of the architecture.

Connectors are hierarchical, tree-like structures with component ports at the leaves. They define sets of *interactions*, based on the attributes of the connected ports [20], which may be either *trigger* (triangles in Fig. 1) or *synchron* (bullets in Fig. 1). If all sub-connectors of a connector are synchrons, then an interaction may be executed by the connector only if each subconnector can contribute. If at least one of the sub-connectors is a trigger, then any interaction consisting of contributions of any set of sub-connectors, *involving at least one of the triggers*, can be executed. For instance, the two ports T.s and C.fail are always synchronised, since they belong to the same binary sub-connector, where they are both synchrons. In particular, this means that whenever the transition $s_1 \xrightarrow{\mathsf{fail}} s_2$ is fired, so is the transition $t_1 \xrightarrow{\mathsf{s},t:=0,z:=\top} t_2$, initialising the timer. The binary connector $T.\mathsf{s}\bullet\!\!-\!\!\bullet C.\mathsf{fail}$ is a sub-connector of a hierarchical connector, where the port B.fail is a trigger. Thus, the above interaction can only happen together with B.fail, forming a ternary interaction. On the contrary, being a trigger, the port B.fail can fire alone, forming a singleton interaction. The composition semantics of BIP systems consists in firing exactly one interaction, enabled through at least one of the top-level connectors, at each execution round.

Finally, *priorities*—defined by a strict partial order on the set of possible interactions—narrow the choice among the enabled interactions at any given round. The default priority is the so-called *maximal progress*, whereby among any two interactions $a \subset b$ (as sets of ports), b has higher priority than a. For example, the port B.fail will never fire alone in a global state, where both T.s and C.fail are enabled.

Application of the Failure Monitor architecture ensures that, whenever a failure is registered in the operand component, the system will be reset, unless a resumption is registered within Max time units (more details in Sect. 5.4).

Figure 2 shows a pNet encoding of the above Failure Monitor architecture. This encoding is structural: each coordinator component is encoded as a pLTS, the operand component—as a hole; connectors of the BIP model are encoded as synchronisation vectors. Each connector that does not involve triggers is trivially encoded by a synchronisation vector comprising the same ports. In order to encode the semantics of the connectors involving triggers, we 1) in the pLTS encoding the coordinator components, add loop transitions to ensure that all ports involved in such connectors are enabled in all states, 2) associate a Boolean value to each of these ports: the original transitions carry the value *true* (e.g. $s_1 \xrightarrow{\mathsf{fail}(\mathit{true})} s_2$), the added loops carry the value *false* (e.g. $s_3 \xrightarrow{\mathsf{fail}(\mathit{false})} s_3$), 3) add to the corresponding synchronisation vectors the Boolean predicate encoding the connector structure. For example, $SV0$ encodes the connector discussed above: the predicate $(b1 = b2) \wedge (b1 \vee b2 \Rightarrow b0)$ means that the "true" transitions C.fail and T.s can only fire together ($b1 = b2$) and whenever one of them fires, B.fail must fire also ($b1 \vee b2 \Rightarrow b0$). This encoding can be systematically obtained for any hierarchical BIP connector [21]. Although,

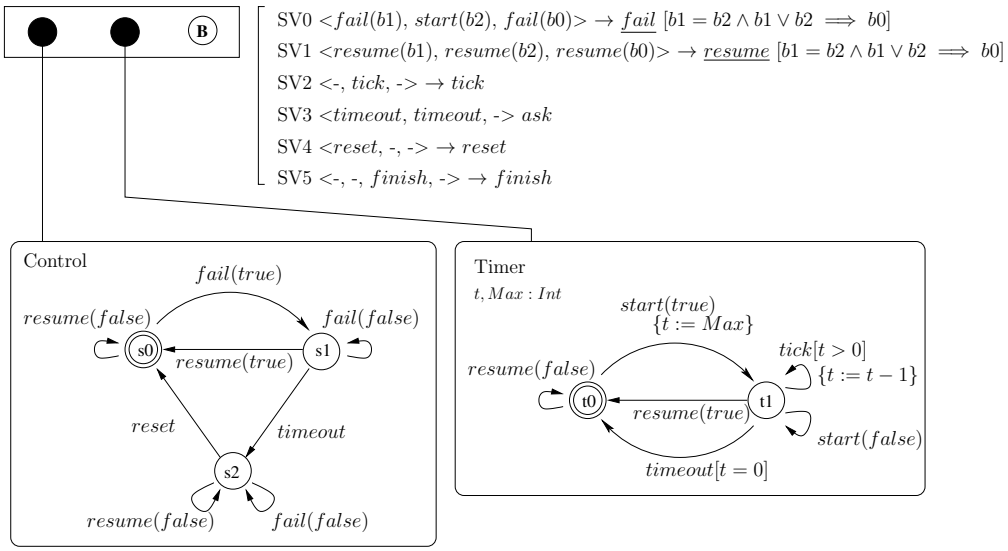

Figure 2: pNet encoding of the Failure Monitor architecture

for the sake of brevity, we omit priorities from the encoding, this can be easily achieved, by introducing additional Boolean variables for relevant ports [14].

4 Operational Semantics for Open pNets

The semantics of open pNets will be defined as an open automaton, that is an automaton where each transition composes transitions of several LTSs with the actions of some holes; the transition occurs if some predicates hold, and can involve a set of state modifications. Each state of an open automaton has a set of *state variables* that can be assigned by incoming transitions. Strictly speaking, the LTSs at the leaves of the open automaton are a restricted form of pLTSs, where labels are parametrised actions, but include no guard nor assignments.

Definition 4 (Open transitions) An *open transition OT* over a set $\langle S_i, s_{0i}, \rightarrow_i \rangle^{i \in I}$ of LTSs, a set J of holes, and a set of states $\mathscr{S}$ is a structure of the form:

$$\frac{\{s_i \xrightarrow{a_i}_i s'_i\}^{i \in I}, \{\xrightarrow{b_j}_j\}^{j \in J}, Pred, Post}{s \xrightarrow{\alpha} s'},$$

where $s, s' \in \mathscr{S}$, the a_i, b_j, α are action expressions, $s_i \xrightarrow{a_i}_i s'_i$ is a transition of the LTS $\langle S_i, s_{0i}, \rightarrow_i \rangle$, b_j is an action of the hole j, and α is the resulting action of OT. $Pred$ is a predicate over the variables of the terms, labels, and states s_i, b_j, s, α. $Post$ is a set of equations that hold *after the open transition*, represented as a substitution $\{x_k \leftarrow e_k\}^{k \in K}$ where x_k are variables of s' and s'_i, whereas e_k are expressions over the other variables of the open transition.

Definition 5 (Open automaton) An *open automaton* is a structure $A = \langle LTS_i^{i \in I}, J, \mathscr{S}, s_0, \mathscr{T} \rangle$ where:

- I and J are sets of indices,

- $LTS_i^{i\in I}$ is a family of LTSs,
- $\mathscr{S}$ is a set of states and $s_0 \in \mathscr{S}$ the initial state,
- $\mathscr{T}$ is a set of open transitions and, for each $t \in \mathscr{T}$, there exist I', J' with $I' \subseteq I$, $J' \subseteq J$, such that t is an open transition over $LTS_i^{i\in I'}$, J', and $\mathscr{S}$.

The *states* and the shape of *predicates* in the transitions of an open automaton representing the semantics of a pNet have the following specific structure.

States of open pNets. A state of an open pNet is a tuple of the states of its leaves (in which we denote tuples in structured states as $\triangleleft \ldots \triangleright$). For any pNet p, let $\overline{Leaves} = \langle\!\langle S_i, s_{i0}, \rightarrow_i \rangle\!\rangle^{i\in L}$ be the set of pLTS at its leaves, then $States(p) = \{\triangleleft s_i^{i\in L} \triangleright \,|\, \forall i \in L. s_i \in S_i\}$. A pLTS being its own single leave: $States(\langle\!\langle S, s_0, \rightarrow \rangle\!\rangle) = \{\triangleleft s \triangleright \,|\, s \in S\}$. The initial state is defined as: $InitState(p) = \triangleleft {s_{i0}}^{i\in L} \triangleright$.

Predicates. Let $\langle\!\langle \overline{pNet}, J, SV_k^{k\in K} \rangle\!\rangle$ be a pNet. Consider a synchronization vector SV_k, for $k \in K$. We build a predicate *MkPred* relating the actions of the involved sub-pNets and the resulting actions. This predicate verifies:

$$MkPred(SV_k, a_i^{i\in I}, b_j^{j\in J}, v) \iff \exists (a'_i)^{i\in I}, (b'_j)^{j\in J}, v'.$$
$$SV_k = (a'_i)^{i\in I}, (b'_j)^{j\in J} \rightarrow v' \wedge \forall i \in I. a_i = a'_i \wedge \forall j \in J. b_j = b'_j \wedge v = v'$$

Example 1 (An open transition) *This transition is generated by application of the vector* `SV0`*, synchronizing the initial actions of pLTSs* `Control` *and* `Timer` *(see Fig. 2), with an action of the hole* `B` *equal to* `fail`*. The local variable* `t` *of the* `Timer` *is assigned to its initial value* `Max`*. A full output of the use-case is provided in [22].*

$$\frac{\{s0 \xrightarrow{fail(true)} s1, t0 \xrightarrow{start(true)} t1\}\{\xrightarrow{hb}\}, hb = fail(true) \wedge v = \underline{fail}, \{t := Max\}}{\triangleleft s0, t0 \triangleright \xrightarrow{v} \triangleleft s1, t1 \triangleright}$$

Structural Semantic Rules: The semantics of pNet in term of open automata has been defined in [8], in the form of 2 structural rules, one for pLTSs, one for pNet nodes. These rules are slightly improved, adding guards in the synchronisation vectors and syntax for universal quantifier in the guards (see [22]).

In [8] we also proved:

Theorem 1 (Finiteness) *Given an open pNet* $pnet = \langle\!\langle \overline{pNet}, \overline{S}, SV_k^{k\in K} \rangle\!\rangle$ *with leaves* $l_i^{i\in I}$ *and holes* $h_j^{j\in J}$*, if the sets I and J are finite, if the synchronisation vectors of all pNets included in pnet are finite, and if* $\forall i \in L. finite(states(l_i))$ *and* l_i *has a finite number of state variables, then its semantics is an open automaton* $\mathscr{T}$ *with finitely many states and transitions.*

Remark that all the elements of such pNets and open automata being symbolic, they can represent many classes of unbounded systems.

5 Generation of Open Automata

In this section we describe the algorithm implementing the pNet semantics, the interaction with the Z3 SMT solver, and we show the result on our example. Under the hypotheses of Theorem 1 this algorithm terminates.

Algorithm 1 Open Automaton Generation

Input: A pNet P (cannot be a hole)

1: Initialize sets $U = \{InitState(P)\}$ and $E = \emptyset$, for unexplored and explored global states, respectively; $L = \emptyset$ for the resulting OTs;
2: **while** $!isEmpty(U)$ **do**
3: Choose S in U; remove S from U, add S to E;
4: $OTs = MakeTransitions(P,S)$;
5: **for** each $OT \in OTs$ **do** Check satisfiability of OT using the SMT solver;
6: **if** $SAT(OT)$ **then**
7: {Add OT to L;
8: Let S' be the target global state of OT
9: **if** $(S' \notin U \cup E)$ **then** Add S' into U; }
10: **end for**
11: **end while**
12: **return** $OA = (InitState(P), L)$;

Algorithm 1 starts with an open pNet, and builds its set of open transitions. Its main loop is a classical residual algorithm: starting from the initial global state, it picks a state in an unexplored set, and computes all possible OTs, adding their target states in the unexplored set, until this set is empty.

The inside loop (*MakeTransitions* method) applies recursively the semantic rules following the structure of the pNet. When applied to a pLTSs at the leaves, we simply take the pLTS transitions of the corresponding local state and use the semantic rule to build the OT[1]. When applied to a pNet node we use two methods, combining and matching, to generate the open transitions in a hierarchical manner, as shown in Alg. 2. This method directly manages the holes of the node, so *MakeTransitions* is never called on a hole.

At the root of the pNet, the predicate of each OT is translated into SMTlib assertions, and checked for satisfiability. The final open automaton includes all satisfiable OTs, and the set of reachable global states.

Combining. The combining method enumerates all the possible behaviours of the subnets as all the possible combinations of their open transitions. Assume that there is a collection of n subnets. We denote $\overline{ot}_i$ the set of open transitions of the i-th subnet (obtained in line 3 of the algorithm); "$-$" means that the subnet is not involved. The combination $\overline{comb}$, a set of n-tuples, is the cartesian product: $\overline{comb} = (\{-\} \cup \overline{ot}_1) \times (\{-\} \cup \overline{ot}_2) \times \cdots \times (\{-\} \cup \overline{ot}_n)$.

[1] We omit detailed presentation of this case for the sake of brevity.

Algorithm 2 MakeTransitions() *for a pNet node*

Input: a pNet node P with subnets $\overline{sn}$ and holes $\overline{hole}$; a global state S.

```
1: Initialize empty list l and set L for sub-transitions and transitions, respectively;
2: for each Subnet in sn̄ do                \\ Recursively apply the semantic rules on the subnets
3:     Store MakeTransitions(Subnet, S) in l;
4: end for
5: comb̄ = Combining(l);
6: for each sv ∈ SV and each comb ∈ comb̄ do
7:     ot = Matching(sv, comb̄, holē);
8:     if (ot is defined) then Store ot in L;    \\ if Matching() succeeds
9: end for
10: return L;
```

Matching. The *Matching* method builds the OTs of a pNet node from those of its subnets. For each synchronisation vector and each possible combination of behaviours of the subnets, as generated by the *Combining* method, it builds the corresponding open transition. Here, we only detail the construction of the predicate. From a synchronization vector $sv = \left((a'_i)^{i\in I}(b'_j)^{j\in J} \to v'\right) \in SV$ and its guard G_k; a tuple of open transitions $C = (ot_i)^{i\in[1,n]} \in \overline{comb}$, such that, for each $i \in [1,n]$, either $ot_i = -$, or the result action of ot_i is a_i; the hole behaviours $Hole = (b_j)^{j\in J}$; and a fresh variable v, representing the result action of the OT under construction, we build the predicate:

$$MkPred(sv, C, \overline{hole}) = \\ (\forall i \in I, a_i = a'_i) \wedge (\forall j \in J, b_j = b'_j) \wedge (v = v') \wedge G_k.$$

Filtering. While matching a vector with a combination tuple, *Matching* tries to filter out some incompatibilities; there may be several reasons why the matching would fail:

- if some subnet is marked as inactive in the vector, and the chosen combination has an active behaviour at this position,
- if some subnet action expression in the vector does not match (by pattern-matching) the corresponding action expression in C,
- or if the whole set of active subnet actions in the vector cannot be matched (by unification) with the corresponding action expressions on the tuple C.

Even when unification succeeds, it is still possible that the resulting predicate would be unsatisfiable, because of some incompatibility involving the guards. In our algorithm, we choose to apply only the simplest filter inside the *Matching* method (before applying the predicate and OT construction). Matching and unification will be checked later, together with the guards collected from synchronisation vectors and from pLTS transtions, using the satisfiability check in the SMT engine.

$$ot = \frac{\{s0 \xrightarrow{fail(true)} s0, t0 \xrightarrow{resume(false)} t0\},\ \{\xrightarrow{hb}\},\ fail(true) = fail(b_1) \wedge resume(false) = start(b_2) \wedge hb = fail(b_0) \wedge v = \underline{fail} \wedge b_1 = b_2 \wedge (b_1 \vee b_2) \Rightarrow b_0,\ \{\}}{\lhd s0, t0 \rhd \xrightarrow{v} \lhd s0, t0 \rhd}$$

Figure 3: One of the unsatisfiable open transitions in the Failure Monitor pNet

5.1 Management of state variable assignments

In a pLTS, there may be several incoming transitions that assign potentially different values to a state variable. To handle such cases, the algorithm manages, for each pLTS state, a list of expressions collected from the assignments of each state variable. For a global state in the open automaton, the set of state variables (which may be used in a transition) is the disjoint union of sets of state variables of the individual pLTS states constituting this global state.

5.2 Pruning the unsatisfiable results

Our matching/filtering strategy builds some transitions where the predicates express incompatible constraints. Even if having an unsatisfiable (symbolic) transition would not be incorrect, we choose to minimize the open automaton (i.e. its number of transitions and states), by checking the predicates for satisfiability. In Fig. 3, we show an unsatisfiable open transition from the open automaton of our running example. It shows the case where the failure controller performs a "fail" action, while the timer executes a "resume". The chosen synchronization vector (SV0 from Fig. 2) does not match with these actions, since it expects T.s. This mismatch is materialised by the predicate fragment "$resume(false) = start(b_2)$".

Checking satisfiability requires some symbolic computation on the action expressions and the predicates, which may depend on the specific theory of the action algebra datatypes. The "Modulo Theory" part of SMT solvers is important here, so that the solver can use specific properties of each action algebra.

5.3 Translation to SMTlib

We check the satisfiability of each open transitions using the SMT solver Z3. In this section, we describe the translation of the algebra presentation, of assignments, and of the predicates.

Our implementation submits satisfiability requests to Z3 using its JAVA API. Here, for readability, we show the Z3 code using its SMT-LIB input language. Note that in the previous sections, the OTs were displayed in a simplified, human readable form. The input and output of our tool, and also the generated SMTlib fragments, are slightly more difficult to read, in particular because of structured names for the fresh variables generated by the algorithm, allowing tracability of the result [22].

Figure 4 shows the translation of the transition of Fig. 3 in the SMTlib syntax. It contains the declaration of the BIP action algebra sorts and constructors, then the declaration of variables, and finally the predicate to be checked, encoded as a set of assertions. Here the behavior of the hole "B" appears as a variable "lhb_B:13:1l", and line 10 shows the constraint that matching with a

```
1  (declare-datatypes () ((Action (fail)(resume)(timeout)(reset)(start)(tick)(ask)
2                 (FUN_Action_Bool (fst Action)(snd Bool))(Synchro (action Action)))))
3  (declare-const |b1:sva_SV0:1:1| Bool)
4  (declare-const |b2:sva_SV0:1:1| Bool)
5  (declare-const |b0:sva_SV0:1:1| Bool)
6  (declare-const |:hb_B:13:1| Action)
7  (declare-const |:ra:1:1| Action)
8  (assert (= (FUN_Action_Bool fail true) (FUN_Action_Bool fail |b1:sva_SV0:1:1|)))
9  (assert (= (FUN_Action_Bool resume false) (FUN_Action_Bool start |b2:sva_SV0:1:1|)))
10 (assert (= |:hb_B:13:1| (FUN_Action_Bool fail |b0:sva_SV0:1:1|)))
11 (assert (= |b1:sva_SV0:1:1| |b2:sva_SV0:1:1|))
12 (assert (or (not (or |b1:sva_SV0:1:1| |b2:sva_SV0:1:1|)) |b0:sva_SV0:1:1|))
13 (assert (= (Synchro fail) |:ra:1:1|))
14 (check-sat)
```

```
unsat
```

Figure 4: The input of the Z3 solver in SMT-LIB language and the output result

synchronisation vector has imposed on the behavior of "B". Here we display also the diagnosis ("sat" or "unsat") generated by Z3.

By lack of space, we have only shown here a use-case with boolean data parameters, so we only require the basic SAT capability of the Z3 engine. It should be clear that more involved examples will require reasonning on predicates of various data types. Z3 provides theories for reasonning on integers, bitstrings, enumerated types, arrays, etc. We also need to deal with some universal quantification on free variables occuring in the guards of synch vectors. For other data structures usual in such models, like records, or recursive structures, the user will have to develop its own theories. Decidability of these theories will condition the completeness of the satisfiability check, and the decidability of model-checking or bisimulation checking.

Production of the SMT-lib code To build the input submitted to Z3 for each OT, we translate the algebra presentation, the predicates and the variable assignments into Z3 (Java-API) calls.

Translation of action algebra presentation. In [22], we define the translation of an algebra presentation into SMTlib declarations (`declare-datatypes` and `declare-fun`). We also formalize a condition of well-formedness of pNets ensures that the generated code is correct, and will not raise runtime errors in the SMT engine. Note that the `declare-datatypes` command comprises both the action constructors from table 1 and also the constant action names from our example. Additionally, we will include axiomatisation of any required functions and predicates of the presentation data-types.

Translation of open transitions. In [22] we formally define all steps of the translation of each open transition, including:

- collect all variables in the transition, and declare them (using `declare-const`)
- check well-formedness and correct typing of expressions,
- translate the predicate into a conjunction of assertions,
- if present in the source state, translate the state-variable assignments into a disjunctive assertion.

This translation ensures that no runtime error will occur in the SMT engine.

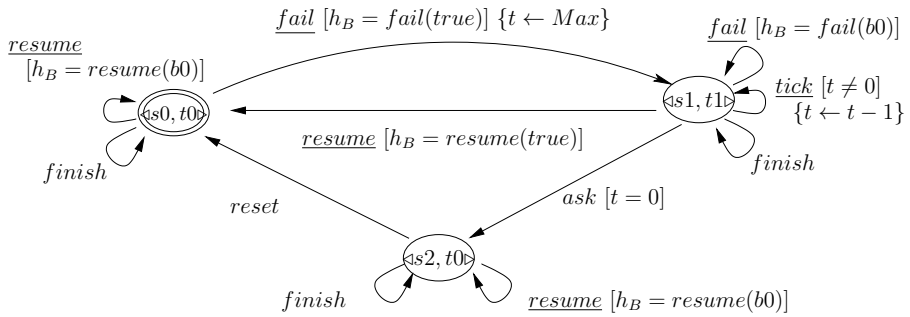

Figure 5: Open Automaton for the Failure Monitor architecture

Figure 4 shows the decomposition of the predicate into a set of asserts, each encoding an elementary equality, inequality, or a guard. The result (sat or unsat) of the final `check-sat` command in the translation () is then decoded.

5.4 Result for the running example

For this example, the tool builds 184 open transitions, whereof 173 are detected unsatisfiable by Z3. The resulting open automaton, with 3 reachable global states (out of the possible 6) and 11 open transitions, is shown in Fig. 5.

To improve the readability of this figure, we used the following conventions: we omit the transitions of the two pLTS, and the set of "working" holes; and we directly write the resulting action as first element of each OT, rather than including it as an equality inside the predicate.

Failure Monitor enforces 1) the safety property *"the system reset never happens, unless asked for by a timeout following a failure"*, formalised in CTL by

$$\varphi \wedge \mathtt{AG}(\mathsf{reset} \rightarrow \varphi), \quad \text{where } \varphi = \mathtt{A}[\neg\mathsf{reset}\ \mathtt{W}\ \mathsf{ask}] \wedge \mathtt{A}[\neg\mathsf{ask}\ \mathtt{W}\ \mathsf{fail}],$$

(`W` being the *weak until* operator) and 2) the liveness property *"a reset will be fired when asked for by a timeout"*: $\mathtt{AG}(\mathsf{ask} \rightarrow \mathtt{AF}\ \mathsf{reset})$. The satisfaction of the safety property could be established by applying symbolic model checking techniques. However, in this example, it is obvious by inspection of the open automaton. The satisfaction of the liveness property relies on the above observation that in the state $\lhd s2, t0 \rhd$ only the reset transition involves C. Therefore, under resonable scheduling assumptions, reset will always be fired.

6 Conclusion and Discussion

The formal definitions and properties of the open pNets model were published in [8]. In this new work we describe an implementation of the model and its semantics construction, including its interaction with the Z3 SMT engine. The implementation has two parts: the first is a finitary algorithm that builds all possible combinations of synchronisations through the pNet hierarchical structure. The result is a so-called *open-automaton*, which transitions contains predicates relating the actions of the pNet holes and controllers. Some of the open transitions obtained at this step, may contain predicates which do not represent any possible concrete instantiations. In the second part of the tool we use the SMT solver Z3 for checking the satisfiability of the predicate in each

open transition. To this end, we encode into Z3 the representations of the action algebra and the predicates before submitting them to the Z3 solver. In this paper, we used a running example, based on a BIP architecture from an earlier nanosatellite case study [19]. This example shows that open-automata-based semantics can be instrumental in verifying the properties enforced by the architectures through an encoding into open pNets. This encoding—which we intend to formalise and prove correct in a separate paper—also opens the way for an extension of BIP architectures with the transfer of data among variables of different components. Indeed, such data transfer can be easily encoded using the predicates associated to synchronisation vectors in open pNets. The encoding of open transitions into SMTlib and the availability of theories can guide the definition of such an extension. Our case-studies show that our encoding successfully identifies the unsatisfiable open transitions and that the resulting automata correctly reflect the expected movements of the encoded process expressions.

Naturally, our next goals after the generation of the open automata will be to model-check logical properties, and to check equivalence of pNets. While model-checking open automata seems easy to define, equivalence checking is more challenging. In [8], we have already found the FH-bisimulation, to be a suitable definition. But weak equivalences, or refinements, will definitely be useful when comparing different pNets with different structure. For bisimulation, we foresee that SMT methods will be the basis for comparison of open transitions.

Scaling up. One important motivation of this work is to attack the complexity of verification of realistic systems by a compositional and parametric approach. Still one may wonder if the price for analysing our symbolic transitions will not make the approach too expensive in term of computing time. We tried a slighly bigger example, assembling 2 Failure controllers. In [22], we show that Z3 can check the satisfiablility of a 90K open transitions in a couple of minutes.

Bibliography

[1] De Simone, R.: Higher-level synchronising devices in MEIJE-SCCS. Theoretical Computer Science **37** (1985) 245–267

[2] Larsen, K.G.: A context dependent equivalence between processes. Theoretical Computer Science **49** (1987) 184–215

[3] Hennessy, M., Lin, H.: Symbolic bisimulations. Theoretical Computer Science **138**(2) (1995) 353–389

[4] Lin, H.: Symbolic transition graph with assignment. In Montanari, U., Sassone, V., eds.: Concur'96. Volume 1119 of LNCS., Springer, Heidelberg (1996) 50–65

[5] Hennessy, M., Rathke, J.: Bisimulations for a calculus of broadcasting systems. Theoretical Computer Science **200**(1-2) (1998) 225–260

[6] Henrio, L., Kulankhina, O., Liu, D., Madelaine, E.: Verifying the correct composition of distributed components: Formalisation and Tool. In: FOCLASA. Number 175 in EPTCS, Rome, Italy (2014)

[7] Henrio, L., Madelaine, E., Zhang, M.: pNets: an Expressive Model for Parameterised Networks of Processes. In: 23rd Euromicro International Conference on Parallel, Distributed, and Network-Based Processing (PDP'15), IEEE (2015)

[8] Henrio, L., Madelaine, E., Zhang, M.: A Theory for the Composition of Concurrent Processes. In: Formal Techniques for Distributed Objects, Components, and Systems (FORTE). Volume LNCS-9688., Heraklion, Greece (2016)

[9] Déharbe, D.: Integration of smt-solvers in b and event-b development environments. Science of Computer Programming **78**(3) (2013) 310–326

[10] Déharbe, D., Fontaine, P., Guyot, Y., Voisin, L.: Integrating smt solvers in rodin. Science of Computer Programming **94** (2014) 130–143

[11] Barrett, C., Conway, C., Deters, M., Hadarean, L., Jovanović, D., King, T., Reynolds, A., Tinelli, C.: Cvc4. In: Computer aided verification, Springer (2011)

[12] Armand, M., Faure, G., Grégoire, B., Keller, C., Théry, L., Werner, B.: A modular integration of sat/smt solvers to coq through proof witnesses. In: International Conference on Certified Programs and Proofs, Springer (2011) 135–150

[13] Basu, A., Bensalem, S., Bozga, M., Combaz, J., Jaber, M., Nguyen, T.H., Sifakis, J.: Rigorous component-based system design using the BIP framework. IEEE Software **28**(3) (2011) 41–48

[14] Baranov, E., Bliudze, S.: Offer semantics: Achieving compositionality, flattening and full expressiveness for the glue operators in BIP. Science of Computer Programming **109**(0) (2015) 2–35

[15] Attie, P., Baranov, E., Bliudze, S., Jaber, M., Sifakis, J.: A general framework for architecture composability. Formal Aspects of Computing **18**(2) (2016) 207–231

[16] Milner, R.: Communication and Concurrency. Int. Series in Computer Science. Prentice-Hall, Englewood Cliffs, New Jersey (1989) SU Fisher Research 511/24.

[17] ISO: Information Processing Systems – Open Systems Interconnection – LOTOS – A Formal Description Technique based on the Temporal Ordering of Observational Behaviour. ISO/IEC 8807, International Organisation for Standardization, Geneva, Switzerland (1989)

[18] Barrett, C., Fontaine, P., Tinelli, C.: The SMT-LIB Standard: Version 2.6. Technical report, Department of Computer Science, The University of Iowa (2017) Available at `www.SMT-LIB.org`.

[19] Mavridou, A., Stachtiari, E., Bliudze, S., Ivanov, A., Katsaros, P., Sifakis, J.: Architecture-based design: A satellite on-board software case study. In: 13th Int. Conf. on Formal Aspects of Component Software (FACS 2016). (2016)

[20] Bliudze, S., Sifakis, J.: The algebra of connectors—Structuring interaction in BIP. IEEE Transactions on Computers **57**(10) (2008) 1315–1330

[21] Bliudze, S., Sifakis, J.: Causal semantics for the algebra of connectors. Formal Methods in System Design **36**(2) (2010) 167–194

[22] Qin, X., Bliudze, S., Madelaine, E., Zhang, M.: Using SMT engine to generate Symbolic Automata – Extended version. Rapport de recherche RR-9177, INRIA (June 2018)