Refined Interfaces for Compositional Verification

Frederic Lang

INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Refined Interfaces for Compositional Verification

Frédéric Lang

N° ????

Octobre 2006

Thème COM

ISSN 0249-6399 ISRN INRIA/RR--????--FR+ENG

Refined Interfaces for Compositional Verification

Frédéric Lang*

Thème COM — Systèmes communicants
Projet VASY

Rapport de recherche n° ???? — Octobre 2006 — 22 pages

Abstract: The compositional verification approach of Graf & Steffen aims at avoiding state space explosion for individual processes of a concurrent system. It relies on interfaces that express the behavioural constraints imposed on each process by synchronization with the other processes, thus preventing the exploration of states and transitions that would not be reachable in the global state space. Krimm & Mounier, and Cheung & Kramer proposed two techniques to generate such interfaces automatically. In this paper, we propose a refined interface generation technique that derives the interface of a process automatically from the examination of (a subset of) concurrent processes. This technique is applicable to formalisms where concurrent processes are composed either using synchronization vectors or process algebra parallel composition operators (including those of CCS, CSP, μCRL, LOTOS, and E-LOTOS). We implemented this approach in the EXP.OPEN 2.0 tool of the CADP toolbox. Several experiments indicate state space reductions by more than two orders of magnitude for the largest processes.

Key-words: Communicating automata, Compositional verification, Concurrency, Enumerative verification, Interface constraints, Formal methods, Parallel composition, Process algebra, Semi-composition

A short version of this report is also available as "*Refined Interfaces for Compositional Verification*", in Elie Najm, Jean-François Pradat-Peyre, and Véronique Viguié Donzeau-Gouge, editors, Proceedings of the 26th IFIP WG 6.1 International Conference on Formal Methods for Networked and Distributed Systems FORTE'2006 (Paris, France), September 26-29, 2006.

* Frederic.Lang@inria.fr

Unit´e de recherche INRIA Rhône-Alpes
655, avenue de l'Europe, 38334 Montbonnot Saint Ismier (France)
T´el´ephone : +33 4 76 61 52 00 — T´el´ecopie +33 4 76 61 52 52

Interfaces raffinées pour la vérification compositionnelle

Résumé : L'approche de vérification compositionnelle de Graf & Steffen a pour but d'éviter l'explosion d'états des processus pris individuellement dans un système concurrent. Elle s'appuie sur des interfaces qui expriment les contraintes comportementales imposées sur chacun des processus par ses synchronisations avec les autres processus, évitant ainsi l'exploration d'états et de transitions qui ne seraient pas atteignables dans l'espace d'états global du système. Krimm & Mounier et Cheung & Kramer ont proposé deux techniques pour générer de telles interfaces automatiquement. Dans ce rapport, nous proposons une technique de génération d'interface raffinée qui dérive automatiquement l'interface d'un processus d'après l'examen (d'un sous ensemble) des processus concurrents. Cette technique est applicable à des formalismes où les processus concurrents sont composés en parallèle, soit en utilisant des vecteurs de synchronisation, soit en utilisant des opérateurs de composition parallèle d'algèbres de processus (incluant ceux de Ccs, Csp, μCrl, Lotos, et E-Lotos). Nous avons implémenté cette approche dans l'outil Exp.Open 2.0 de la boîte à outils Cadp. Plusieurs expérimentations indiquent des réductions d'espace d'états de plus de deux ordres de grandeur pour les processus les plus gros.

Mots-clés : algèbre de processus, automates communicants, composition parallèle, contraintes d'interface, méthodes formelles, parallélisme, semi-composition, vérification compositionnelle, vérification énumérative

1 Introduction

Enumerative verification is a popular technique that consists in exploring and checking reachable states and transitions of a concurrent system. It is confronted with the *state explosion* problem, which occurs when the number of states grows exponentially as the number of concurrent processes increases. To avoid or reduce state explosion, various approaches have been proposed, among which symbolic verification, on-the-fly verification, partial order reductions, symmetries, data-flow analysis, and compositional verification. This paper deals with the latter approach, which assumes that the concurrent system under study can be expressed as a collection of communicating sequential processes, the behaviours of which are modeled as finite state machines or LTSs (*Labelled Transition Systems*). The sequential processes are composed in parallel, either in a flat or hierarchical manner.

In its simplest forms [10, 28, 32, 38, 33, 34, 36, 31], compositional verification (also called incremental reduction [32], incremental reachability analysis [33, 34], compositional state space generation [36], or inductive compression [31]) consists in replacing each sequential process by an *abstraction*, simpler than the original process but still preserving the properties to be verified on the whole system. Quite often, abstracting a process is done by minimizing its corresponding LTS modulo an appropriate equivalence or preorder relation (e.g., a bisimulation relation, such as strong, branching, or observational equivalence). If the system has a hierarchical structure, minimization can also be applied at every intermediate level in the hierarchy. Although this simple form of compositional verification has been applied successfully to some complex systems (e.g., [11, 5] in the case of the LOTOS language [22]), it may be counter-productive in some other cases: generating the LTS of each process separately may lead to state explosion, whereas the generation of the whole system of concurrent processes might succeed if processes constrain each other when composed in parallel. Indeed, there may be many states of a process that, although useful in a general environment, are useless (i.e., never explored) in a particular environment.

This issue has been addressed by enhanced compositional verification approaches [19, 7, 37, 8, 9, 18, 26, 6, 16], which permit the generation of the LTS of an individual process by taking into account *interface constraints* (also known as *environment constraints* or *context constraints*). These constraints express the behavioural restrictions imposed on the considered process by synchronization with its neighbour processes. Taking into account the environment of a process permits local elimination of states and transitions unreachable in the LTS of the whole system.

In general, interface constraints are expressed in the form of an LTS simply called *interface*. There exist two approaches to restrict the behaviour of a process w.r.t. an interface. In the first one, the process is composed in parallel with the interface, which must have been transformed beforehand so that the composition does not affect the global behaviour of the system (a property known as *context transparency*) [6, 7, 8, 9]. This approach is supported in the framework of CSP by the TRACTA tool [16]. In the second approach, the process is constrained using a specific *semi-composition* operator [19, 18, 26], which cuts the process states and transitions that cannot be reached when considering the traces of the interface as the only possible interactions between the process and its environment. This approach is

supported in the framework of LOTOS by the PROJECTOR [26] and SVL [12] tools of CADP (*Construction and Analysis of Distributed Processes*) [13] and was used in the verification of an industrial protocol [35].

Interfaces can be either written by the user (and possibly checked automatically [26]), or generated automatically. Although automated generation has the neat advantage to relieve users from the burden of calculating appropriate constraints, existing automated interface generation techniques undergo two main limitations: first, they are specific to a given composition operator and thus not directly applicable in the framework of concurrent languages featuring different and/or more general operators; second, as already observed in [7], they may fail to capture effective interface constraints due to deficiencies in their analysis of synchronizations between processes[1].

In this paper, we propose to generate interfaces automatically using a new technique that relies on a translation of the system into an intermediate concurrent model, named *network of* LTS*s*, which describes the synchronization between processes in a flat manner. This intermediate representation permits the derivation of effective interface constraints imposed on a given process by a set of its neighbour processes automatically, independently of the hierarchy of processes and of the nature of the composition operators. This permits combination of constraints induced by distant processes, and improvement of the accuracy of interfaces by exploiting more precisely the synchronizations between processes. For this reason, we qualify as *refined* the interfaces generated using this technique.

As regards practical aspects, we implemented refined interface generation in the EXP.OPEN 2.0 tool for on-the-fly verification of networks of LTSs [27] of CADP. Interfaces can be generated automatically from systems made of LTSs composed using operators taken from several languages (CCS [29], CSP [4], μCRL [21], LOTOS [22], the E-LOTOS international standard [24], and general concurrent specification formalisms). In the framework of LOTOS specifications, the SVL scripting language was also extended to facilitate the combined use of the various CADP tools involved to use refined interfaces in a compositional verification task. For behavioural restriction, we rely on PROJECTOR and its semi-composition operator, which is general enough to be applicable in the framework of the above concurrent languages, although originally designed for LOTOS.

Using a flat intermediate concurrent model such as networks of LTSs is not new, as most model-checkers start by flattening the process hierarchy, for instance generating an intermediate Petri net [14] in the case of LOTOS, *Linear Process Equations* in the case of μCRL [20], or using a *supercombinator*-based compilation mechanism called *supercompilation* [17] in the case of CSP. The model we use in this paper is close to MEC *synchronization vectors* [1] and FC2 *synchronization networks* [3]. The originality of our work resides in both the treatment we make on the intermediate model to generate interfaces, and the effective use of this model to handle many different operators in a compositional verification setting.

This paper focuses on communication by *rendez-vous* between processes which run asynchronously (i.e., at independent speeds). It naturally generalizes to communication through bounded buffers if buffers are represented as finite processes communicating by *rendez-vous*

[1] See in particular Examples 2 and 3, Section 3 of this paper.

with the rest of the system[2]. The current approach can be used to constrain such buffers in the same way as any process. Approaches to constrain processes communicating through buffers that are not bounded *a priori* (i.e., the bound of each buffer, if any, is not known statically but determined at execution time) have been proposed [25] but are out of the scope of this paper.

The paper is organized as follows: Section 2 presents the technical background. Section 3 recalls semi-composition and discusses the limitations of existing interface generation methods. Section 4 defines refined interface generation, which improves over existing interface generation methods. Section 5 describes the implementation of refined interface generation in CADP. Section 6 presents some experimental results. Section 7 finally concludes.

2 Technical Background

Definition 1 (Vectors) A *vector* of *length* n over a set S is an element of S^n, written $\boldsymbol{t}$ or $(t_1, \ldots, t_n)$. For $i \in 1..n$, $\boldsymbol{t}[i]$ denotes the ith element t_i of $\boldsymbol{t}$, and $\boldsymbol{t}[i \leftarrow t'_i]$ represents a copy of $\boldsymbol{t}$ where $\boldsymbol{t}[i]$ is replaced by t'_i. Given $t \in S$, we write t^n the vector of length n such that $(\forall i \in 1..n)\ t^n[i] = t$. Given $I \subseteq 1..n$, the *projection* $\boldsymbol{t}_{\downarrow I}$ is defined by: $\boldsymbol{t}_{\downarrow I} = (\boldsymbol{t}[k_1], \ldots, \boldsymbol{t}[k_m])$ where $\{k_i \mid i \in 1..m\} = I$ and $(\forall i < j)\ k_i < k_j$.

Definition 2 (Labelled Transition System) Let $\mathcal{A}$ be a set of symbols called *observable actions*, and $\tau \notin \mathcal{A}$ the *unobservable action*. Given $A \subseteq \mathcal{A}$, we write A_τ the set $A \cup \{\tau\}$. An LTS is a quadruple $S = (Q, A, T, q_0)$, where Q is the set of *states*, $A \subseteq \mathcal{A}$ — also written $act(S)$ — is the set of *observable actions*, $T \subseteq Q \times A_\tau \times Q$ is the *transition relation*, and $q_0 \in Q$ is the *initial state*. As usual, we may write $q_1 \xrightarrow{a}_T q_2$ (or $q_1 \xrightarrow{a} q_2$ when T is clear from the context) instead of $(q_1, a, q_2) \in T$. A *trace* of S is a sequence of actions $a_1 \ldots a_{n \geq 0} \in (A_\tau)^n$, such that $(\exists q_1, \ldots, q_n \in Q)\ (\forall i \in 0..n-1)\ q_i \xrightarrow{a_{i+1}}_T q_{i+1}$ (note that the sequence starts in the initial state q_0 of S). An *observable trace* is a trace in which all occurrences of τ have been removed. We write $\mathcal{L}(S)$ the set of observable traces of S. An action $a \in A$ is *reachable* if there is a trace containing a. A state $q \in Q$ is *reachable* if there exists a trace such that $q_n = q$. A transition $(q_1, a, q_2) \in T$ is *reachable* if q_1 is reachable. Two LTSs S_1, S_2 are *equal*, written $S_1 = S_2$, if and only if they have the same initial states and reachable transitions.

3 Semi-Composition

Semi-composition [26] (implemented in the PROJECTOR tool of CADP) permits restriction of the behaviour of a process *on-the-fly* by taking into account interface constraints, usually derived from its environment. Since semi-composition was designed in the framework of

[2] See `http://www.inrialpes.fr/vasy/cadp/case-studies` which references more than 80 case studies in various application domains, many of which use bounded buffers.

LOTOS, its definition is tightly related to the following LOTOS-like parallel composition and hiding operators.

Definition 3 (Parallel Composition, Hiding) Let $S_i = (Q_i, A_i, T_i, q_{0i})$ $(i = 1, 2)$ be two LTSs, and $A \subseteq \mathcal{A}$. The *parallel composition* "$S_1 \parallel_A S_2$" models the concurrent execution of S_1 and S_2 with forced synchronization on A. It is defined as the LTS $(Q, A_1 \cup A_2, T, (q_{01}, q_{02}))$, where Q and T are the smallest sets satisfying both $(q_{01}, q_{02}) \in Q$ and the following properties:

$$\frac{(q_1, q_2) \in Q,\ q_1 \xrightarrow{a}_{T_1} q_1',\ q_2 \xrightarrow{a}_{T_2} q_2',\ a \in A}{(q_1', q_2') \in Q,\ (q_1, q_2) \xrightarrow{a}_{T} (q_1', q_2')}$$

$$\frac{(q_1, q_2) \in Q,\ q_1 \xrightarrow{a}_{T_1} q_1',\ a \notin A}{(q_1', q_2) \in Q,\ (q_1, q_2) \xrightarrow{a}_{T} (q_1', q_2)} \qquad \frac{(q_1, q_2) \in Q,\ q_2 \xrightarrow{a}_{T_1} q_2',\ a \notin A}{(q_1, q_2') \in Q,\ (q_1, q_2) \xrightarrow{a}_{T} (q_1, q_2')}$$

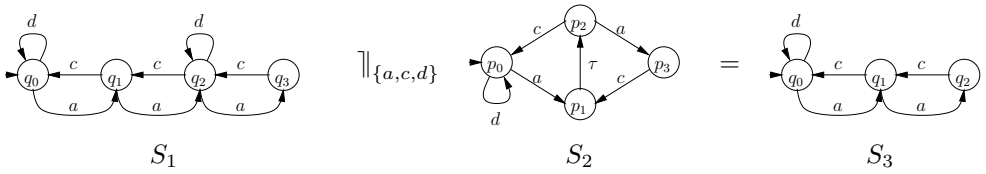
Note that, by construction, the states belonging to Q are reachable. A state p of S_1 (respectively S_2) is said *reachable* in $S_1 \parallel_A S_2$ if there is a state (p, q) (resp. (q, p)) in $S_1 \parallel_A S_2$. Similarly, a transition $p \xrightarrow{a} p'$ of S_1 (respectively S_2) is said *reachable* in $S_1 \parallel_A S_2$ if there is a transition $(p, q) \xrightarrow{a} (p', q')$ (resp. $(q, p) \xrightarrow{a} (q', p')$) in $S_1 \parallel_A S_2$. The expression "hide A in S_1" denotes the LTS $(Q_1, A_1 \setminus A, T_1', q_{01})$, where T_1' is defined as follows:

$$\frac{q \xrightarrow{a}_{T_1} q',\ a \in A}{q \xrightarrow{\tau}_{T_1'} q'} \qquad \frac{q \xrightarrow{a}_{T_1} q',\ a \notin A}{q \xrightarrow{a}_{T_1'} q'}$$

Semi-composition takes as input two LTSs S_1, S_2 and a set of actions A, and returns the LTS which contains exactly the states and transitions of S_1 that are reachable in $S_1 \parallel_A S_2$.

Definition 4 (Semi-Composition) Let $S_i = (Q_i, A_i, T_i, q_{0i})$ $(i = 1, 2)$ be two LTSs, $A \subseteq \mathcal{A}$, and $(Q', A', T', q_0') = S_1 \parallel_A S_2$. The *semi-composition of S_1 and S_2*, written "$S_1 \rceil\!\rceil_A S_2$", is the LTS (Q, A_1, T, q_{01}), where $Q = \{p \mid (p, q) \in Q'\}$ and $T = T_1 \cap \{(p_1, a, p_2) \mid (p_1, q_1) \xrightarrow{a}_{T'} (p_2, q_2)\}$. A is called the *synchronization set* and the pair (A, S_2) is called the *interface*[3]. We say that an action $a \in A_1$ is *controlled* by the interface (A, S_2) if $a \in A$.

Example 1 The following holds:



S_1 S_2 S_3

State q_3 and transitions $q_2 \xrightarrow{d} q_2, q_2 \xrightarrow{a} q_3$, and $q_3 \xrightarrow{c} q_2$ do not belong to S_3 because they are not reachable in $S_1 \parallel_{\{a,c,d\}} S_2$.

[3]This definition of semi-composition is simpler but equivalent to that given in [26].

Three properties of semi-composition are essential to ensure its practicability:

- Semi-composition is a state space reduction, since the sets of states and transitions of $S_1 \rceil\!\rceil_A S_2$ are by definition subsets of S_1. The worst case is when $\mathcal{L}(\text{hide } (\mathcal{A} \backslash A) \text{ in } S_1) \subseteq \mathcal{L}(\text{hide } (\mathcal{A} \setminus A) \text{ in } S_2)$, yielding $S_1 \rceil\!\rceil_A S_2 = S_1$.
- $(S_1 \rceil\!\rceil_A S_2) \parallel_A S_2 = S_1 \parallel_A S_2$. Therefore semi-composition can be used to reduce S_1 given its environment S_2 by removing the unreachable states and transitions, without losing any temporal property of the system $S_1 \parallel_A S_2$. Note that, unlike Cheung & Kramer's approach, which requires that the interface be context transparent — and thus be transformed into a deterministic LTS using a well-known but expensive algorithm — no restriction is made here on the shape of S_2.
- $S_1 \rceil\!\rceil_A S_2 = S_1 \rceil\!\rceil_A S_2'$ if $\mathcal{L}(\text{hide } (\mathcal{A} \setminus A) \text{ in } S_2) = \mathcal{L}(\text{hide } (\mathcal{A} \setminus A) \text{ in } S_2')$. Therefore, reductions of the interface can be achieved by first hiding uncontrolled actions and then minimizing the LTS modulo a relation preserving observable traces (e.g., *safety equivalence* [2]), which permits reduction of the number of states to explore while calculating semi-composition. Safety minimization is less expensive than determinization and, unlike determinization which can induce a dramatic growth of the LTS, yields an LTS that contains fewer states than the input. Minimization of the interface is not mandatory but important to reduce the cost of semi-composition, the complexity of which is the same as parallel composition, hence sensitive to the size of its operands.

In practice, the equation $S_1 \parallel_A S_2 = (S_1 \rceil\!\rceil_A S_2) \parallel_A S_2$ is not sufficient to compute interfaces in the case of systems consisting of more than two LTSs: it may happen that S_2 does not constrain S_1 but that a more distant LTS in the environment of S_1 does. Krimm & Mounier proposed a method to compute an exact interface in the framework of more general systems of communicating LTSs built upon parallel composition and action hiding. Given two LTSs S_1 and S_2 in such a system, this method permits to synthesize a synchronization set A such that S_1 can be replaced by $S_1 \rceil\!\rceil_A S_2$ without changing the global LTS of the system. It is defined inductively, based on the following semi-composition laws:

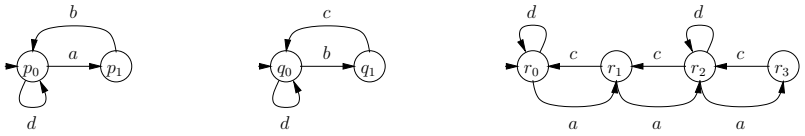
$$S_1 \parallel_A S_2 = (S_1 \rceil\!\rceil_A S_2) \parallel_A S_2 \quad (1)$$

$$(S_1 \parallel_{A_1} S_3) \parallel_{A_2} S_2 = ((S_1 \rceil\!\rceil_B S_2) \parallel_{A_1} S_3) \parallel_{A_2} S_2 \quad (2)$$
$$\text{where } B = A_2 \cap (A_1 \cup (act(S_1) \setminus act(S_3)))$$

$$(\text{hide } A_1 \text{ in } S_1) \parallel_{A_2} S_2 = (\text{hide } A_1 \text{ in } (S_1 \rceil\!\rceil_{A_2 \setminus A_1} S_2)) \parallel_{A_2} S_2 \quad (3)$$

Unfortunately, the interface (A, S_2) built using Krimm & Mounier's method generally does not give the best account of environment constraints, as illustrated by the following two examples.

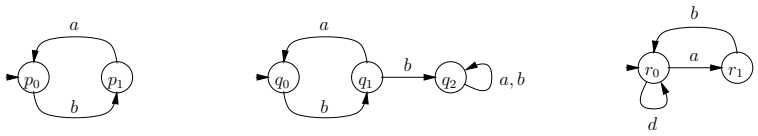
Example 2 Let $E = S_1 \parallel_{\{a,b,d\}} (S_2 \parallel_{\{c,d\}} S_3)$ with S_1, S_2, and S_3 as follows:



S_1 S_2 S_3

According to the semi-composition laws, S_3 can be replaced in E either by $S_3 \rceil\!\rceil_{\{a,d\}} S_1$, or by $S_3 \rceil\!\rceil_{\{c,d\}} S_2$, but both expressions result in S_3 itself, due to the fact that $\mathcal{L}(\text{hide } (\mathcal{A} \setminus \{a,d\}) \text{ in } S_1) = \{a,d\}^*$ and $\mathcal{L}(\text{hide } (\mathcal{A} \setminus \{c,d\}) \text{ in } S_2) = \{c,d\}^*$. Yet, one can see that actions a and c are executed with some alternation in E, due to the mandatory synchronization on b between S_1 and S_2. As a consequence, state r_3 is not reachable in E. To capture such a constraint, it should be possible to build an interface that takes simultaneously into account the constraints induced by both S_1 and S_2, even though there is no sub-expression of E containing S_1 and S_2 only. This is not possible with Krimm & Mounier's method[4].

Example 3 Let $E = S_1 \parallel_{\{a,b\}} (S_2 \parallel_{\{a\}} S_3)$ with S_1, S_2, and S_3 as follows:



S_1 S_2 S_3

According to the semi-composition laws, S_2 can be replaced by $S_2 \rceil\!\rceil_{\{a\}} S_1$, but this expression yields S_2 itself, due to the fact that $\mathcal{L}(\text{hide } (\mathcal{A} \setminus \{a\}) \text{ in } S_1) = a^*$. Yet, it is clear from S_1 and the synchronizations in E that state q_2 of S_2 is unreachable in E, as two successive b actions cannot be fired without an a in between. A better interface should permit to take into account the environment constraints due to synchronizations on b, even though every b of S_1 does not necessarily synchronize with a b of S_2. Unfortunately, this is not possible using the Krimm & Mounier's method[5].

In the sequel, we propose to generate interface constraints automatically in a way that palliates these limitations.

[4]This limitation holds similarly for Cheung & Kramer's method, as mentioned in [7].

[5]Cheung & Kramer do not provide a solution to this issue as their method relies on a CSP-like parallel composition operator whose semantics states that synchronization on b is mandatory between all processes containing b in their action set.

4 Refined Interface Generation

Refined interface generation is a new method that permits the computation of an interface capturing the constraints imposed on a given process P in a concurrent system by one or several processes of its environment. This interface can then be semi-composed with P on-the-fly, so as to restrict P's behaviour.

As regards the model of concurrency on which we establish our results, we use the following network model named "*network of* LTSs", in which the composition hierarchy is completely flattened. The network of LTSs model is more general than the parallel composition operator defined in the previous section, and the parallel composition, renaming, hiding and cutting operators from many process algebras can be translated into networks of LTSs [27]. Networks of LTSs thus make our work non-specific to a particular process algebra and permit an easier way of reasoning about the synchronization structure of systems.

Definition 5 (Network of LTSs) Let $\bullet \notin \mathcal{A}_\tau$ be a special symbol denoting that a particular LTS has no role in a given synchronization. A *synchronization rule* is a pair $(\boldsymbol{t}, a)$, where $\boldsymbol{t}$ is a vector over $\mathcal{A}_\tau \cup \{\bullet\}$ (called a *synchronization vector*) and $a \in \mathcal{A}_\tau$. The components $\boldsymbol{t}$ and a are called respectively the left- and right-hand sides of the synchronization rule. A *network of* LTS*s* (or simply *network*) N of *dimension* $n > 0$ is a pair $(\boldsymbol{S}, V)$ where $\boldsymbol{S}$ is a vector of LTSs of length n and V is a set of synchronization rules, whose left-hand sides are all of length n. Each left-hand side $\boldsymbol{t}$ expresses a synchronization constraint on $\boldsymbol{S}$, all components $\boldsymbol{S}[i]$ where $\boldsymbol{t}[i] \neq \bullet$ having to take a transition labeled respectively $\boldsymbol{t}[i]$ altogether so that a transition labeled with the corresponding right-hand side a be generated in the product. More formally, let $\boldsymbol{S}[i] = (Q_i, A_i, T_i, q_{0i})$ $(i \in 1..n)$. To $N = (\boldsymbol{S}, V)$ corresponds an LTS $(Q, A, T, \boldsymbol{q_0})$, written $sem(N)$ or $sem(\boldsymbol{S}, V)$, such that $A = \{a \mid (\boldsymbol{t}, a) \in V\}$, $\boldsymbol{q_0} = (q_{01}, \ldots, q_{0n})$, and Q and T are the smallest sets satisfying both $\boldsymbol{q_0} \in Q$ and:

$$\frac{\boldsymbol{q} \in Q,\ (\boldsymbol{t}, a) \in V,\ (\forall i \in 1..n)\ (\boldsymbol{t}[i] = \bullet \wedge \boldsymbol{q'}[i] = \boldsymbol{q}[i]) \vee \boldsymbol{q}[i] \xrightarrow{\boldsymbol{t}[i]}_{T_i} \boldsymbol{q'}[i]}{\boldsymbol{q'} \in Q,\ (\boldsymbol{q}, a, \boldsymbol{q'}) \in T}$$

Note that, by construction, the states that belong to Q are reachable. Synchronization rules must obey the following *admissibility* properties, which forbid cutting, synchronizations and renaming of τ transitions and therefore ensure that safety equivalence and stronger relations (e.g., observational, branching, and strong equivalences) are congruences for networks of LTS [27]:

$$((\exists i \in 1..n)\ \tau \text{ is reachable in } \boldsymbol{S}[i]) \implies (\exists (\boldsymbol{t}, \tau) \in V)\ \boldsymbol{t}[i] = \tau$$

$$(\forall (\boldsymbol{t}, a) \in V)\ ((\exists i \in 1..n)\ \boldsymbol{t}[i] = \tau) \implies (a = \tau \wedge (\forall j \in 1..n \setminus \{i\})\ \boldsymbol{t}[j] = \bullet)$$

Example 4 Systems of communicating LTSs built upon various operators can be translated into networks of LTSs. As an example, given S_1 and S_2, the parallel composition $(S_1 \parallel_A S_2)$ can be translated into $((S_1, S_2), V_{sync} \cup V_{async})$, where:

$$\begin{aligned} V_{sync} &= \{((a, a), a) \mid a \in act(S_1) \cap act(S_2) \cap A\} \\ V_{async} &= \{((a, \bullet), a) \mid a \in act(S_1)_\tau \setminus A\} \cup \{((\bullet, a), a) \mid a \in act(S_2)_\tau \setminus A\} \end{aligned}$$

Note that sets of synchronization rules are more general than functions because nondeterminism is allowed, i.e., they may contain two synchronization rules $(\boldsymbol{t}, a)$ and $(\boldsymbol{t}, b)$ with identical synchronization vector $\boldsymbol{t}$; if the synchronization w.r.t. $\boldsymbol{t}$ is possible in a given state, then two transitions labelled respectively a and b result. Note also that there may exist synchronization rules of the form $(\bullet^n, a)$, which yield a looping transition labelled a in every reachable state.

Given a network $N = (\boldsymbol{S}, V)$ and an LTS $\boldsymbol{S}[k]$ in this network, we address the problem of computing automatically an interface of the form $(\mathcal{A}, C)$ that will permit reduction of $\boldsymbol{S}[k]$ by taking into account its interactions with a subset $\{\boldsymbol{S}[i] \mid i \in I\}$ $(k \notin I)$ of LTSs in its environment. The goal is to permit the replacement of LTS $\boldsymbol{S}[k]$ by LTS $\boldsymbol{S}[k] \parallel_{\mathcal{A}} C$ in N without affecting the LTS of the global system. To this aim, we define the following refined interface generation procedure, whose inputs are N, k, and I. The refined interface generated consists of a product of the LTSs $\boldsymbol{S}[i]$ $(i \in I)$, synchronized by synchronization rules derived systematically from the synchronization rules of N, each rule $(\boldsymbol{t}, a)$ being transformed into a rule $(\boldsymbol{t}_{\downarrow I}, \boldsymbol{t}[k])$ if $\boldsymbol{t}[k] \neq \bullet$, or $(\boldsymbol{t}_{\downarrow I}, \tau)$ otherwise. Therefore, whenever a transition $\boldsymbol{q} \xrightarrow{a} \boldsymbol{q}'$ can be fired in $sem(N)$ using a synchronization rule $(\boldsymbol{t}, a)$ with $\boldsymbol{t}[k] \neq \bullet$, then the participating transition $\boldsymbol{q}[k] \xrightarrow{\boldsymbol{t}[k]} \boldsymbol{q}'[k]$ of $\boldsymbol{S}[k]$ is also a transition of $\boldsymbol{S}[k] \parallel_{\mathcal{A}} C$. Conversely, transitions of $\boldsymbol{S}[k]$ that cannot participate in any mandatory synchronization with C (i.e., the $\boldsymbol{S}[i]$'s) are eliminated by the semi-composition $\boldsymbol{S}[k] \parallel_{\mathcal{A}} C$.

Definition 6 (Refined Interface Generation) Let $\varphi : A_\tau \cup \{\bullet\} \rightarrow A_\tau$, defined by $\varphi(\bullet) = \tau$ and $(\forall a \in A_\tau)\ \varphi(a) = a$. Let $N = (\boldsymbol{S}, V)$ be a network of dimension n, I a set of indices such that $\emptyset \subset I \subset 1..n$, and k an index such that $k \in 1..n \setminus I$. The *refined interface* of $\boldsymbol{S}[k]$ capturing constraints induced by $\{\boldsymbol{S}[i] \mid i \in I\}$, written $refint(N, k, I)$, is the interface $(\mathcal{A}, sem(\boldsymbol{S}_{\downarrow I}, V'))$, where $V' = \{(\boldsymbol{t}_{\downarrow I}, \varphi(\boldsymbol{t}[k])) \mid (\boldsymbol{t}, a) \in V\}$.

Example 5 Consider the network N displayed on the left below, with arbitrary LTSs $S_1, \ldots, S_4$. The refined interface of S_1 capturing constraints induced by S_3 and S_4, written $refint(N, 1, \{3, 4\})$, is the LTS corresponding to the network displayed on the right below. Note the projection on S_3 and S_4, and observe that the right-hand sides of synchronization rules in the result are the elements of column S_1, where $\bullet$ is renamed into τ.

$$refint\left(\left(\begin{array}{c}(S_1, S_2, S_3, S_4),\\ \left\{\begin{array}{l}((a_1, a_2, a_3, a_4),\ a),\\ ((\ \bullet,\ b_2,\ b_3,\ \bullet),\ b),\\ ((\ c_1,\ c_2,\ \bullet,\ \bullet),\ c)\end{array}\right\}\end{array}\right), 1, \{3, 4\}\right) = sem\left(\begin{array}{c}(S_3, S_4),\\ \left\{\begin{array}{l}((a_3, a_4), a_1),\\ ((\ b_3,\ \bullet),\ \tau),\\ ((\ \bullet,\ \bullet),\ c_1)\end{array}\right\}\end{array}\right)$$

The following theorem states that, in an arbitrary network N, any interface $refint(N, k, I)$ can be used to restrict $\boldsymbol{S}[k]$ using semi-composition because the LTS of N and the LTS of N in which $\boldsymbol{S}[k]$ is replaced by its restriction are equal.

Theorem 1 Let $N = (\boldsymbol{S}, V)$ be a network of dimension n, I such that $\emptyset \subset I \subset 1..n$, $k \in 1..n \setminus I$, and $(A, C) = refint(N, k, I)$. If $\boldsymbol{S}' = \boldsymbol{S}\big[k \leftarrow (\boldsymbol{S}[k] \parallel_A C)\big]$ then $sem(\boldsymbol{S}, V) = sem(\boldsymbol{S}', V)$.

Proof. Since $\boldsymbol{S}[k] \rceil\!\rceil_A C$ is a sub-LTS of $\boldsymbol{S}[k]$ by definition of semi-composition, it follows that $sem(\boldsymbol{S}', V)$ is a sub-LTS of $sem(\boldsymbol{S}, V)$. We show that, conversely, $sem(\boldsymbol{S}, V)$ is a sub-LTS of $sem(\boldsymbol{S}', V)$. To this aim, we consider an arbitrary state $\boldsymbol{q}$ reachable in $sem(\boldsymbol{S}, V)$. In a first step we assume that $\boldsymbol{q}_{\downarrow I}$ is reachable in C, $(\boldsymbol{q}[k], \boldsymbol{q}_{\downarrow I})$ is reachable in $\boldsymbol{S}[k] \parallel_{\mathcal{A}} C$, $\boldsymbol{q}[k]$ is reachable in $\boldsymbol{S}'[k]$, $\boldsymbol{q}$ is reachable in $sem(\boldsymbol{S}', V)$ and given a transition $\boldsymbol{q} \xrightarrow{a} \boldsymbol{q}'$ of $sem(\boldsymbol{S}, V)$ induced by a vector $(\boldsymbol{t}, a)$, we show simultaneously that (1) $\boldsymbol{q}'_{\downarrow I}$ is reachable in C, (2) $(\boldsymbol{q}'[k], \boldsymbol{q}'_{\downarrow I})$ is reachable in $\boldsymbol{S}[k] \parallel_{\mathcal{A}} C$, which implies that $\boldsymbol{q}'[k]$ is reachable in $\boldsymbol{S}'[k]$, and (3) $\boldsymbol{q} \xrightarrow{a} \boldsymbol{q}'$ is a transition of $sem(\boldsymbol{S}', V)$, which implies that $\boldsymbol{q}'$ is reachable in $sem(\boldsymbol{S}', V)$. We consider two cases:

- If $\boldsymbol{t}[k] = \bullet$ then by definition $\boldsymbol{q}[k] = \boldsymbol{q}'[k]$ and property (3) is obvious. In addition, by definition of *refint*, the transition $\boldsymbol{q}_{\downarrow I} \xrightarrow{\tau} \boldsymbol{q}'_{\downarrow I}$ belongs to C, which implies properties (1) and (2).

- If $\boldsymbol{t}[k] \neq \bullet$ then by hypothesis $\boldsymbol{q}[k] \xrightarrow{\boldsymbol{t}[k]} \boldsymbol{q}'[k]$ belongs to $\boldsymbol{S}[k]$ and $\boldsymbol{q}_{\downarrow I} \xrightarrow{\boldsymbol{t}[k]} \boldsymbol{q}'_{\downarrow I}$ belongs to C by definition of *refint*, which implies property (1). Therefore, $(\boldsymbol{q}[k], \boldsymbol{q}_{\downarrow I}) \xrightarrow{\boldsymbol{t}[k]} (\boldsymbol{q}'[k], \boldsymbol{q}'_{\downarrow I})$ belongs to $\boldsymbol{S}[k] \parallel_{\mathcal{A}} C$, which implies property (2). By definition of semi-composition, $\boldsymbol{q}[k] \xrightarrow{\boldsymbol{t}[k]} \boldsymbol{q}'[k]$ belongs to $\boldsymbol{S}'[k]$, which implies property (3).

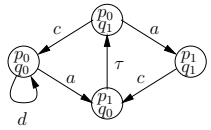
In a second step, given $\boldsymbol{q}_0$ the initial state of $sem(\boldsymbol{S}, V)$, we observe that $\boldsymbol{q}_{0\downarrow I}$, $(\boldsymbol{q}_0[k], \boldsymbol{q}_{0\downarrow I})$, $\boldsymbol{q}_0[k]$, and $\boldsymbol{q}_0$ are the initial states of, respectively, C, $\boldsymbol{S}[k] \parallel_{\mathcal{A}} C$, $\boldsymbol{S}'[k]$, and $sem(\boldsymbol{S}', V)$. Given a state $\boldsymbol{q}$ reachable in $sem(\boldsymbol{S}, V)$, an induction using properties (1), (2), and (3) shows that $\boldsymbol{q}_{\downarrow I}$, $(\boldsymbol{q}[k], \boldsymbol{q}_{\downarrow I})$, $\boldsymbol{q}[k]$, and $\boldsymbol{q}$ are reachable in, respectively, C, $\boldsymbol{S}[k] \parallel_{\mathcal{A}} C$, $\boldsymbol{S}'[k]$, and $sem(\boldsymbol{S}', V)$. Therefore, every transition of $sem(\boldsymbol{S}, V)$ is also a transition of $sem(\boldsymbol{S}', V)$, which implies that $sem(\boldsymbol{S}, V)$ and $sem(\boldsymbol{S}', V)$ are equal. □

The following examples show that refined interfaces solve the issues raised in Examples 2 and 3 of Section 3.

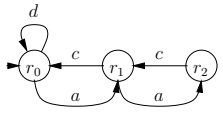
Example 6 [back to Example 2 page 7] Expression $E = S_1 \parallel_{\{a,b,d\}} (S_2 \parallel_{\{c,d\}} S_3)$ defined in Example 2 can be translated into the network N displayed below. S_3 may be restricted using a refined interface $(\mathcal{A}, sem(N')) = \mathit{refint}(N, 3, \{1, 2\})$ that takes simultaneously both S_1 and S_2 into account, where N' and $sem(N')$ are displayed below. $S_3 \rceil\!\rceil_{\mathcal{A}} sem(N')$, also displayed below, reduces S_3 by eliminating the unreachable state r_3 and transitions $r_2 \xrightarrow{a} r_3, r_3 \xrightarrow{c} r_2$,

and $r_2 \xrightarrow{d} r_2$.

$$N = \left(\begin{array}{c} (S_1, S_2, S_3), \\ \left\{ \begin{array}{l} ((\ a,\ \bullet,\ a), a), \\ ((\ b,\ b,\ \bullet), b), \\ ((\ \bullet,\ c,\ c), c), \\ ((\ d,\ d,\ d), d) \end{array} \right\} \end{array} \right) \qquad N' = \left(\begin{array}{c} (S_1, S_2), \\ \left\{ \begin{array}{l} ((\ a,\ \bullet), a), \\ ((\ b,\ b), \tau), \\ ((\ \bullet,\ c), c), \\ ((\ d,\ d), d) \end{array} \right\} \end{array} \right)$$



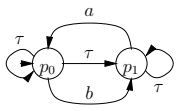
$sem(N')$



$S_3 \rceil\!\rceil_{\mathcal{A}}\ sem(N')$

Example 7 [back to Example 3 page 8] Expression $E = S_1 \parallel_{\{a,b\}} (S_2 \parallel_{\{a\}} S_3)$ defined in Example 3 can be translated into the network N displayed below. S_2 may be restricted using a refined interface $(\mathcal{A}, sem(N')) = refint(N, 2, \{1\})$ that takes S_1 into account, where N' and $sem(N')$ are displayed below. In practice, $sem(N')$ can be minimized modulo safety equivalence, yielding an LTS with 2 states and 3 transitions. $S_2 \rceil\!\rceil_{\mathcal{A}}\ sem(N')$ is isomorphic to S_1.

$$N = \left(\begin{array}{c} (S_1, S_2, S_3), \\ \left\{ \begin{array}{l} ((\ a,\ a,\ a), a), \\ ((\ b,\ b,\ \bullet), b), \\ ((\ b,\ \bullet,\ b), b), \\ ((\ \bullet,\ \bullet,\ d), d) \end{array} \right\} \end{array} \right) \quad N' = \left(\begin{array}{c} (S_1), \\ \left\{ \begin{array}{l} ((\ a\ ), a), \\ ((\ b\ ), b), \\ ((\ b\ ), \tau), \\ ((\ \bullet\ ), \tau) \end{array} \right\} \end{array} \right) \qquad$$



$sem(N')$

This example shows that without using more LTSs from the environment of S_2 than in Example 3, but simply by taking a better account of the synchronization structure of the system, the *refint* operation permits refinement of the interface with respect to that obtained using equation (2), turning the set of observable traces of the interface from a^* with b uncontrolled in Example 3 to $a^* + b + (ba^+)^*$ in the current example. The latter set of traces does not contain any trace with two consecutive b's, thus disabling the transition $q_1 \xrightarrow{b} q_2$ in S_2 and making state q_2 and transitions $q_2 \xrightarrow{a} q_2$, $q_2 \xrightarrow{b} q_2$ also unreachable.

The *refint* operation may create synchronization rules of the form $(\bullet^n, a)$, which induce a self-looping transition labelled a in each state of the interface (see for instance the last synchronization rule of the right-hand side network in Example 5 and the last synchronization rule of network N' in Example 7, which induces the τ-loops in states p_0 and p_1). Some of these synchronization rules can be eliminated as follows:

- Every synchronization rule of the form $(\bullet^n, \tau)$ can merely be removed. Indeed, for all $\boldsymbol{S}$ and V, $\mathcal{L}(sem(\boldsymbol{S}, V \cup (\bullet^n, \tau))) = \mathcal{L}(sem(\boldsymbol{S}, V))$.

- Every synchronization rule of the form $(\bullet^n, a)$ where $a \neq \tau$ can be removed if the set of synchronization rules does not contain another rule with the same action a as right-hand side. Indeed, for all $\boldsymbol{S}$, S', A, and V in which a does not occur as a right-hand side, $S' \rrbracket_A \ sem(\boldsymbol{S}, V \cup (\bullet^n, a)) = S' \rrbracket_{A \setminus a} \ sem(\boldsymbol{S}, V)$. Eliminating this rule transforms the synchronization set of the interface from A into $A \setminus a$.

Algorithmically, refined interface generation has the same complexity as the synchronization product of the LTSs taken into account in the environment. In practice, the cost of computing the interface can be reduced by minimizing the individual LTSs participating in the interface modulo safety equivalence, which is correct due to the above mentioned congruence property of safety equivalence. In addition, well-known partial order reductions preserving observable traces can be used to further reduce interfaces on-the-fly during their construction.

So far, refined interface generation required that each (high-level) process of the concurrent system under verification was replaced by its LTS, which apparently contradicts the claim that refined interfaces can be used to restrict processes on-the-fly. However, it is clear from Definition 6 that the states and transitions of LTS $\boldsymbol{S}[k]$ (corresponding to the process to restrict) are not needed for interface generation. In practice, only the observable actions of $\boldsymbol{S}[k]$ are needed to compute the synchronization rules of the network from higher level operators as in Example 4. To do so, $\boldsymbol{S}[k]$ can be replaced by an abstraction consisting of an arbitrary (and much smaller) LTS containing the same set of actions. In fact, the method remains correct if the abstraction contains a superset of $\boldsymbol{S}[k]$'s actions, although the reduction obtained on $\boldsymbol{S}[k]$ by semi-composition generally increases while the set of actions of the abstraction gets closer to the exact set of actions of $\boldsymbol{S}[k]$.

In practice, users must provide such an abstraction "by hand", which is not hard as it suffices to examine the gates (or channels) occurring in the process specification and the types of their data, and to enumerate actions of this type appropriately. If the abstraction provided by the user lacks some action of $\boldsymbol{S}[k]$, then the generated interface might be wrong, but this is detected automatically during the compositional verification task as explained in [26]. Calculating this abstraction automatically from source code or from an internal representation of processes would not present any difficulty.

5 Implementation in the CADP Toolbox

Our method was implemented in CADP (*Construction and Analysis of Distributed Processes*) [13], a popular toolbox for protocol engineering. Refined interface generation is implemented as an option (**-interface**) of the EXP.OPEN 2.0 tool [27] for on-the-fly verification of products of communicating LTSs, which can be combined using the following operators:

- standard parallel composition, action cutting, action hiding, and action renaming from CCS, CSP, LOTOS, and μCRL;

- networks of LTSs and generalized parallel composition from E-LOTOS, which includes n-ary parallel composition, "n among m" parallel composition, and parallel composition with synchronization interfaces [15];
- generalized forms of action hiding, action renaming, and transition cutting, where actions can be defined using regular expressions.

EXP.OPEN 2.0 also implements several partial order reductions, one of which can be used to partially reduce the interface on-the-fly while preserving its observable traces (**-weaktrace** option).

To simplify the use of refined interfaces in the more specific framework of LOTOS descriptions, we have also extended the SVL scripting language [12] with a new operator, named "`refined abstraction`", which can be used in the context of any parallel composition expression. As an example, given a LOTOS file `"file.lotos"` defining the system "`(P |[A, C]| Q) |[A, B]| R`", where `P`, `Q`, and `R` are LOTOS processes, one may write the following SVL script:

```
% DEFAULT_LOTOS_FILE="file.lotos"
"file.bcg" = root leaf strong reduction of
((refined abstraction Q, R using "act.bcg" of P) |[A, C]| Q) |[A, B]| R
```

This script computes the LTS corresponding to the system by first restricting `P` on-the-fly w.r.t. the constraints induced by `Q` and `R`, using the LTS `"act.bcg"` as the abstraction of `P`. To this aim, `Q` and `R` are first minimized modulo safety equivalence and an interface generated automatically using EXP.OPEN 2.0. Once the LTSs corresponding to processes `P` (restricted using the refined interface), `Q`, and `R` have been generated, the "`root leaf strong reduction`" operation minimizes them modulo strong bisimulation, and then minimizes their product once they have been composed in parallel. The result is stored in `"file.bcg"`.

6 Applications

We applied refined interfaces to three case studies. The first one is a LOTOS description written by J. Romijn [30] of the HAVi (*Home Audio-Video*) asynchronous leader election protocol[6], which consists of seven concurrent processes named `BUSRESET`, `DCM1`, `DCM2`, `CMM1`, `CMM2`, `MS1`, and `MS2`. Given a LOTOS process `ABS_DCM1` containing the actions of `DCM1`, we made the following experiments:

E1 Generation of `DCM1` without interface.

E2 Generation of `DCM1` using an interface consisting of the LTS of the sub-system including `CMM1, CMM2, MS1`, and `MS2`, and of a synchronization set computed as defined by Krimm & Mounier's semi-composition laws.

[6] See `ftp://ftp.inrialpes.fr/pub/vasy/demos/demo_27`

Exp.	Interface				DCM1	
	generated		minimized (safety)		generated	
	states	trans.	states	trans.	states	trans.
E1	0	0	0	0	**404,477**	**3,025,842**
E2	3,904	42,697	3	37	**365,923**	**2,514,848**
E3	704	7,145	4	45	**17,199**	**73,130**
E4	2,328	14,158	52	613	**645**	**2,020**

Exp.	Interface generation		Interface safety minimization		DCM1 generation		DCM1 strong minimization		Total time	Max memory
E1	0 s	0 Mb	0 s	0 Mb	69.9 s	11 Mb	30.0 s	54 Mb	99.9 s	54 Mb
E2	3.0 s	1.9 Mb	37.4 s	10.5 Mb	115.4 s	26 Mb	26.3 s	46 Mb	182.1 s	46 Mb
E3	3.0 s	1.4 Mb	2.3 s	5.9 Mb	6.1 s	4 Mb	0.7 s	1.9 Mb	12.1 s	5.9 Mb
E4	3.2 s	1.4 Mb	5.1 s	8.5 Mb	2.3 s	3 Mb	0.1 s	1.9 Mb	10.7 s	8.5 Mb

Figure 1: LTS sizes, computation time and memory consumption for experiments E1-E4.

E3 Generation of DCM1 using a refined interface capturing the constraints induced by CMM1, CMM2, MS1, and MS2.

E4 Same as E3, capturing also the constraints induced by BUSRESET and DCM2.

Figure 1 contains two tables. The first table indicates for each experiment E1 to E4 the size of the interface before and after safety minimization, and the size of DCM1 restricted by the interface (if any). The second table indicates computation times and memory consumption for the different operations. They show that refined interfaces permit state space reductions by more than two orders of magnitude (from 404,477 states reachable in a general environment down to 645 states reachable in an environment that takes an account of all processes — experiment E4), while globally reducing verification time by a factor of almost 10 and peak memory consumption by a factor of up to 9.

Experiments E2 and E3 take an account of the same processes to restrict DCM1, the difference being that E2 uses Krimm & Mounier's method and E3 the *refint* operation to compute the interface. Figure 1 thus shows that *refint* yields an LTS with more than 20 times fewer states and 35 times fewer transitions than Krimm & Mounier's method, while the execution time and peak memory consumption are reduced by factors of 15 and 8 respectively. Note that Krimm & Mounier's method does not permit the computation of an interface that takes an account of all processes in a way analogous to E4, because the processes in the environment of DCM1 belong to different sub-expressions.

Second, we considered an ODP (*Open Distributed Processing*) trader [23], an E-LOTOS model of which was presented in [15][7]. An ODP trader is an agent that registers services that can be provided by distant servers, receives service requests from distant clients, and provides to the requesting clients the address of a server that can furnish the requested service. The

[7]See ftp://ftp.inrialpes.fr/pub/vasy/demos/demo_37

client and server are then able to exchange the service directly without communicating with the trader anymore. Note that the trader is a central component in the ODP model in the sense that the ability of two agents to communicate is initiated by the trader. Such central components generally have large state spaces, especially in compositional verification settings where their LTS have to be generated outside of any context.

In our experiment, the components (trader, clients and services) are described in LOTOS and the synchronization structure describing their interactions in EXP.OPEN 2.0 using the "n among m" E-LOTOS parallel composition operator to model the dynamicity of object exchanges. In this example, the ODP trader executes in an environment consisting of 4 objects and 5 services. A refined interface is generated automatically from this environment to restrict the LTS corresponding to the trader, which is thus limited to 256 states instead of 1 million otherwise.

At last, we studied a standard cache coherency protocol for multiprocessor architectures, which consists of a remote directory process named `REMOTE_DIRECTORY` and several agent processes named `AGENT_1` to `AGENT_5`, accessing the directory concurrently[8]. In a configuration with 5 agents, refined interface generation has allowed us to reduce the size of the LTS corresponding to the remote directory from 1 million states and 40 million transitions downto less than 60 states, whereas Krimm & Mounier's method did not yield any state space reduction. In the Figures 2, 3, and 4, we report the results of the following experiments F1 to F3:

F1 (Direct generation.) The global LTS corresponding to the specification is generated at once (using the CÆSAR.OPEN and GENERATOR tools) and minimized modulo strong bisimulation (using BCG_MIN).

F2 (Compositional verification without interface.) The LTS corresponding to each process is generated (using CÆSAR.OPEN and GENERATOR) and minimized modulo strong bisimulation (using BCG_MIN). The resulting LTSs are then composed, and the LTS corresponding to the composition is generated (using EXP.OPEN 2.0 and GENERATOR). At last, the resulting LTS is minimized modulo strong bisimulation (using BCG_MIN).

F3 (Refined interface using `AGENT_1`.) The LTS corresponding to the `AGENT_`i processes are generated (using CÆSAR.OPEN and GENERATOR), as well as a chaos LTS containing the actions of `REMOTE_DIRECTORY`. Those LTSs are minimized modulo safety equivalence (using ALDEBARAN), composed in an expression with the same architecture as the specification, and used to extract a refined interface capturing constraints imposed by `AGENT_1` on `REMOTE_DIRECTORY` (using EXP.OPEN 2.0 with `-interface` option and GENERATOR). The interface constrained LTS corresponding to `REMOTE_DIRECTORY` is then generated (using CÆSAR.OPEN and PROJECTOR), minimized modulo strong bisimulation (using BCG_MIN), and composed with the LTSs corresponding to the `AGENT_`i, also minimized modulo strong bisimulation. The LTS corresponding to this

[8] See `ftp://ftp.inrialpes.fr/pub/vasy/demos/demo_28`

	Experiment F2				Experiment F3			
	Generation		Minimization		Generation		Minimization	
n	states	trans.	states	trans.	states	trans.	states	trans.
3	2,881	51,196	180	3,180	25	102	10	41
4	51,451	1,423,867	2,058	56,742	36	188	12	62
5	1,032,193	40,871,972	28,672	1,132,544	49	310	14	87
6	explosion				64	474	16	116
7	explosion				81	686	18	149

Figure 2: Size of the Lts corresponding to `REMOTE_DIRECTORY` in experiments F2 and F3.

	Experiment F1		Experiment F3		Minimal (strong)	
n	states	trans.	states	trans.	states	trans.
3	6,322	18,444	816	2,639	545	1,787
4	149,924	540,956	5,113	21,761	3,172	13,477
5	4,209,604	17,957,714	30,514	160,885	17,227	90,087
6	explosion		176,667	1,110,211	89,434	556,101
7	explosion		1,001,876	7,306,051	449,593	3,242,227

Figure 3: Size of the global Lts corresponding to the protocol in experiments F1 and F3 and size once minimized modulo strong bisimulation.

	Experiment F1		Experiment F2		Experiment F3	
n	Time	Memory	Time	Memory	Time	Memory
3	4 s	2 MB	14 s	2 MB	17 s	3 MB
4	27 s	12 MB	1 min	25 MB	22 s	3 MB
5	26 min 44 s	349 MB	34 min 15 s	660 MB	35 s	5 MB
6	explosion		explosion		1 min 52 s	25 MB
7	explosion		explosion		12 min 29 s	155 MB

Figure 4: Total CPU time used and maximal memory consumed in experiments F1 to F3.

composition is generated (using Exp.Open 2.0 and Generator) and the resulting Lts is minimized modulo strong bisimulation.

The figures show that only experiment F3 (using refined interfaces) allows us to generate the global state space corresponding to the cache coherency protocol specification with 7 agents, while both other approaches lead to state explosion as soon as $n = 6$.

7 Conclusion

Compositional verification in which the behaviours of concurrent processes are restricted using interface constraints is an effective method to avoid the state explosion that may occur when the state space of a process is generated out of its environment. This paper alleviates the lack of efficient methods to synthesize constraints automatically, by proposing a method based on the analysis of the synchronizations between concurrent processes.

Compared to prior work [7, 9, 26, 6], our method performs a finer analysis of synchronization constraints: our implementation in the Exp.Open 2.0 tool of Cadp exhibits more than two orders of magnitude better state space reductions on an industrial case study studied by Romijn [30]. Moreover, it provides a systematic way of using the semi-composition operator of Krimm & Mounier [26] (which is implemented in the Projector tool of Cadp) in the framework of languages whose composition operators are not limited to Lotos parallel composition and hiding; indeed, both synchronization vectors and a large number of parallel composition operators are supported, including those of Ccs, Csp, Lotos, μCrl, and E-Lotos. Alternatively, we believe that we can also use parallel composition instead of semi-composition as advocated by Cheung & Kramer [7, 9, 6]; indeed the interfaces generated for semi-composition can be transformed into "context-transparent" interfaces using the algorithm given in [7].

Acknowledgements

The author thanks the anonymous referees, and Hubert Garavel, Radu Mateescu, Gwen Salaün, and Wendelin Serwe from the Vasy team at Inria Rhône-Alpes for useful comments on this paper and on earlier versions of this paper.

References

[1] André Arnold. MEC: A System for Constructing and Analysing Transition Systems. In Joseph Sifakis, editor, *Proceedings of the 1st Workshop on Automatic Verification Methods for Finite State Systems (Grenoble, France)*, volume 407 of *Lecture Notes in Computer Science*, pages 117–132. Springer Verlag, June 1989.

[2] Ahmed Bouajjani, Jean-Claude Fernandez, Susanne Graf, Carlos Rodríguez, and Joseph Sifakis. Safety for Branching Time Semantics. In *Proceedings of 18th ICALP*. Springer Verlag, July 1991.

[3] Amar Bouali, Annie Ressouche, Valérie Roy, and Robert de Simone. The Fc2Tools set: a Toolset for the Verification of Concurrent Systems. In Rajeev Alur and Thomas A. Henzinger, editors, *Proceedings of the 8th Conference on Computer-Aided Verification (New Brunswick, New Jersey, USA)*, volume 1102 of *Lecture Notes in Computer Science*. Springer Verlag, August 1996.

[4] S. D. Brookes, C. A. R. Hoare, and A. W. Roscoe. A Theory of Communicating Sequential Processes. *Journal of the ACM*, 31(3):560–599, July 1984.

[5] Ghassan Chehaibar, Hubert Garavel, Laurent Mounier, Nadia Tawbi, and Ferruccio Zulian. Specification and Verification of the PowerScale Bus Arbitration Protocol: An Industrial Experiment with LOTOS. In Reinhard Gotzhein and Jan Bredereke, editors, *Proceedings of the Joint International Conference on Formal Description Techniques for Distributed Systems and Communication Protocols, and Protocol Specification, Testing, and Verification FORTE/PSTV'96 (Kaiserslautern, Germany)*, pages 435–450. IFIP, Chapman & Hall, October 1996. Full version available as INRIA Research Report RR-2958.

[6] K. H. Cheung. *Compositional Analysis of Complex Distributed Systems*. PhD thesis, Department of Computer Science, Hong Kong University of Science and Technology, Hong Kong, 1998.

[7] S. C. Cheung and J. Kramer. Enhancing Compositional Reachability Analysis with Context Constraints. In *Proceedings of the 1st ACM SIGSOFT International Symposium on the Foundations of Software Engineering (Los Angeles, CA, USA)*, pages 115–125. ACM Press, December 1993.

[8] S. C. Cheung and J. Kramer. Compositional Reachability Analysis of Finite-State Distributed Systems with User-Specified Constraints. In *Proceedings of the 3rd ACM SIGSOFT International Symposium on the Foundations of Software Engineering (Washington, DC, USA)*, pages 140–150. ACM Press, October 1995.

[9] S. C. Cheung and J. Kramer. Context Constraints for Compositional Reachability. *ACM Transactions on Software Engineering Methodology TOSEM*, 5(4):334–377, October 1996.

[10] Jean-Claude Fernandez. *ALDEBARAN : un système de vérification par réduction de processus communicants*. Thèse de Doctorat, Université Joseph Fourier (Grenoble), May 1988.

[11] Jean-Claude Fernandez, Hubert Garavel, Laurent Mounier, Anne Rasse, Carlos Rodríguez, and Joseph Sifakis. A Toolbox for the Verification of LOTOS Programs. In Lori A. Clarke, editor, *Proceedings of the 14th International Conference on Software Engineering ICSE'14 (Melbourne, Australia)*, pages 246–259. ACM, May 1992.

[12] Hubert Garavel and Frédéric Lang. SVL: a Scripting Language for Compositional Verification. In Myungchul Kim, Byoungmoon Chin, Sungwon Kang, and Danhyung Lee, editors, *Proceedings of the 21st IFIP WG 6.1 International Conference on Formal Techniques for Networked and Distributed Systems FORTE'2001 (Cheju Island, Korea)*, pages 377–392. IFIP, Kluwer Academic Publishers, August 2001. Full version available as INRIA Research Report RR-4223.

[13] Hubert Garavel, Frédéric Lang, and Radu Mateescu. An Overview of CADP 2001. *European Association for Software Science and Technology (EASST) Newsletter*, 4:13–24, August 2002. Also available as INRIA Technical Report RT-0254 (December 2001).

[14] Hubert Garavel and Joseph Sifakis. Compilation and Verification of LOTOS Specifications. In L. Logrippo, R. L. Probert, and H. Ural, editors, *Proceedings of the 10th International Symposium on Protocol Specification, Testing and Verification (Ottawa, Canada)*, pages 379–394. IFIP, North-Holland, June 1990.

[15] Hubert Garavel and Mihaela Sighireanu. A Graphical Parallel Composition Operator for Process Algebras. In Jianping Wu, Qiang Gao, and Samuel T. Chanson, editors, *Proceedings of the Joint International Conference on Formal Description Techniques for Distributed Systems and Communication Protocols, and Protocol Specification, Testing, and Verification FORTE/PSTV'99 (Beijing, China)*, pages 185–202. IFIP, Kluwer Academic Publishers, October 1999.

[16] D. Giannakopoulou. *Model Checking for Concurrent Software Architectures*. PhD thesis, Imperial College of Science, Technology and Medicine — University of London — Department of Computer Science, January 1999.

[17] Michael Goldsmith. Operational Semantics for Fun and Profit. In Ali E. Abdallah, Cliff B. Jones, and Jeff W. Sanders, editors, *Proceedings of the Symposium on the Occasion of 25 Years of* CSP *(London, UK)*, volume 3525 of *Lecture Notes in Computer Science*. Springer, 2005.

[18] S. Graf, B. Steffen, and G. Lüttgen. Compositional Minimisation of Finite State Systems using Interface Specifications. *Formal Aspects of Computation*, 8(5):607–616, September 1996.

[19] Susanne Graf and Bernhard Steffen. Compositional Minimization of Finite State Systems. In R. P. Kurshan and E. M. Clarke, editors, *Proceedings of the 2nd Workshop on Computer-Aided Verification (Rutgers, New Jersey, USA)*, volume 531 of *Lecture Notes in Computer Science*, pages 186–196. Springer Verlag, June 1990.

[20] Jan Friso Groote and Michel Reniers. *Algebraic Process Verification*. In S.A. Smolka J.A. Bergstra, A. Ponse, editor, *Handbook of Process Algebra*, chapter 17, pages 1151–1208. North-Holland, 2001.

[21] J.F. Groote and A. Ponse. Syntax and semantics of μ-CRL. In *Proceedings of* Algebra of Communicating Processes, Workshops in Computing, 1995.

[22] ISO/IEC. LOTOS — A Formal Description Technique Based on the Temporal Ordering of Observational Behaviour. International Standard 8807, International Organization for Standardization — Information Processing Systems — Open Systems Interconnection, Genève, September 1989.

[23] ISO/IEC. Open Distributed Processing – Reference Model. International Standard 10746, International Organization for Standardization — Information Processing Systems, Genève, 1995.

[24] ISO/IEC. Enhancements to LOTOS (E-LOTOS). International Standard 15437:2001, International Organization for Standardization — Information Technology, Genève, September 2001.

[25] Jean-Pierre Krimm. *Application des ordres partiels à la génération compositionnelle de systèmes asynchrones*. Thèse de Doctorat, Université Joseph Fourier, Grenoble, December 2000.

[26] Jean-Pierre Krimm and Laurent Mounier. Compositional State Space Generation from LOTOS Programs. In Ed Brinksma, editor, *Proceedings of TACAS'97 Tools and Algorithms for the Construction and Analysis of Systems (University of Twente, Enschede, The Netherlands)*, volume 1217 of *Lecture Notes in Computer Science*, Berlin, April 1997. Springer Verlag. Extended version with proofs available as Research Report VERIMAG RR97-01.

[27] Frédéric Lang. EXP.OPEN 2.0: A Flexible Tool Integrating Partial Order, Compositional, and On-the-fly Verification Methods. In Jaco van de Pol, Judi Romijn, and Graeme Smith, editors, *Proceedings of the 5th International Conference on Integrated Formal Methods IFM'2005 (Eindhoven, The Netherlands)*, Lecture Notes in Computer Science. Springer Verlag, November 2005. Full version available as INRIA Research Report RR-5673.

[28] J. Malhotra, S. A. Smolka, A. Giacalone, and R. Shapiro. A Tool for Hierarchical Design and Simulation of Concurrent Systems. In *Proceedings of the BCS-FACS Workshop on Specification and Verification of Concurrent Systems (Stirling, Scotland)*, pages 140–152, Swinton, UK, July 1988. British Computer Society.

[29] Robin Milner. *Communication and Concurrency*. Prentice-Hall, 1989.

[30] Judi Romijn. Model Checking the HAVi Leader Election Protocol. Technical Report SEN-R9915, CWI, Amsterdam, The Netherlands, June 1999. submitted to Formal Methods in System Design.

[31] A.W. Roscoe. *The Theory and Practice of Concurrency*. Prentice Hall, 1998.

[32] K. K. Sabnani, A. M. Lapone, and M. U. Uyar. An Algorithmic Procedure for Checking Safety Properties of Protocols. *IEEE Transactions on Communications*, 37(9):940–948, September 1989.

[33] K. C. Tai and V. Koppol. Hierarchy-Based Incremental Reachability Analysis of Communication Protocols. In *Proceedings of the IEEE International Conference on Network Protocols (San Francisco, CA)*, pages 318–325, Piscataway, NJ, October 1993. IEEE Press.

[34] K. C. Tai and V. Koppol. An Incremental Approach to Reachability Analysis of Distributed Programs. In *Proceedings of the 7th International Workshop on Software Specification and Design (Los Angeles, CA)*, pages 141–150, Piscataway, NJ, December 1993. IEEE Press.

[35] Frédéric Tronel, Frédéric Lang, and Hubert Garavel. Compositional Verification Using CADP of the ScalAgent Deployment Protocol for Software Components. In Uwe Nestmann and Perdita Stevens, editors, *Proceedings of the 6th IFIP International Conference on Formal Methods for Open Object-based Distributed Systems FMOODS'2003 (Paris, France)*, volume 2884 of *Lecture Notes in Computer Science*, pages 244–260. Springer Verlag, November 2003. Full version available as INRIA Research Report RR-5012.

[36] Antti Valmari. Compositional State Space Generation. In *Proceedings of Advances in Petri Nets*, volume 674 of *Lecture Notes in Computer Science*, pages 427–457. Springer Verlag, 1993.

[37] W. J. Yeh. *Controlling State Explosion in Reachability Analysis*. PhD thesis, Software Engineering Research Center (SERC) Laboratory, Purdue University, December 1993. Technical Report SERC-TR-147-P.

[38] W. J. Yeh and M. Young. Compositional Reachability Analysis Using Process Algebra. In *Proceedings of the ACM SIGSOFT Symposium on Testing, Analysis, and Verification (SIGSOFT'91, Victoria, British Columbia, Canada)*, pages 49–59, New York, NY, October 1991. ACM Press.

Unit´e de recherche INRIA Rhône-Alpes
655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)

Unit´e de recherche INRIA Futurs : Parc Club Orsay Universit´e - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)
Unit´e de recherche INRIA Lorraine : LORIA, Technop^ole de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unit´e de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unit´e de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)
Unit´e de recherche INRIA Sophia Antipolis : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399