IF: An Intermediate Representation and Validation Environment for Timed Asynchronous Systems

Marius Bozga[1], Jean-Claude Fernandez[2], Lucian Ghirvu[1]*, Susanne Graf[1], Jean-Pierre Krimm[1], and Laurent Mounier[1]

[1] VERIMAG*** Centre Equation, 2 avenue de Vignate, F-38610 Gières
Marius.Bozga@imag.fr
[2] LSR/IMAG, BP 82, F-38402 Saint Martin d'Hères Cedex
Jean-Claude.Fernandez@imag.fr

Abstract. Formal Description Techniques (FDT), such as LOTOS or SDL are at the base of a technology for the specification and the validation of telecommunication systems. Due to the availability of commercial tools, these formalisms are now being widely used in the industrial community. Alternatively, a number of quite efficient verification tools have been developed by the research community. But, most of these tools are based on simple ad hoc formalisms and the gap between them and real FDT restricts their use at industrial scale.
This context motivated the development of an intermediate representation called IF which is presented in the paper. IF has a simple syntactic structure, but allows to express in a convenient way most useful concepts needed for the specification of timed asynchronous systems. The benefits of using IF are multiples. First, it is general enough to handle significant subsets of most FDTs, and in particular a translation from SDL to IF is already implemented. Being built upon a mathematically sound model (extended timed automata) it allows to properly evaluate different semantics for FDTs, in particular with respect to time considerations. Finally, IF can serve as a basis for interconnecting various tools into a unified validation framework. Several levels of IF program representations are already available via well defined APIs and allow to connect tools ranging from static analyzers to model-checkers.

keywords: asynchrony, timed systems, timed automata, model-checking, static analysis

1 Introduction

Formal Description Techniques, such as LOTOS [ISO88] or SDL [IT94] and related formalisms such as MSC and TTCN are at the base of a technology for the specification and the validation of telecommunication systems. Due to the availability of commercial tools, mainly for editing, code generation and testing, and the fact that these formalisms are promoted by ITU and other international standardization bodies, these formalisms are now being widely used in the community of telecommunication systems.

* Work partially supported by Région Rhône-Alpes, France
*** Verimag is Research Laboratory of CNRS, Université Joseph Fourier and Institut National Polytechnique of Grenoble

There are also increasing needs for description and validation tools covering as many aspects of system development as possible. This is the reason why the commercial editing tools contain also some verification facilities. Unfortunately, these verification facilities are often quite restricted and the tools are "closed" in the sense that there are only limited possibilities to interface them with others. On the other hand, a number of quite efficient verification tools have been developed by the research community, but they are in general based on ad hoc input formalisms and the gap between them and real FDT restricts their use at an industrial scale. Even if these tools are in general less closed than commercial ones, they have rarely well-defined interfaces. For example, a lot of developments were made around the Spin verification tool [Hol91], but they are based on the availability of the source code and not on *a priori* defined interfaces.

A different approach was followed within CADP [FGK+96], a toolbox for the verification of LOTOS specifications. It was conceived right from the beginning as an open platform for interfacing different algorithms and provides several well-defined and documented interfaces (Application Programming Interfaces, API for short). The initial motivation for the work presented here was the fact that SDL becomes a more and more popular formalism in the telecommunication community, and that we wanted to adapt CADP to deal also with SDL specifications. Since the intermediate program level formalisms used within CADP are not appropriate for SDL specifications, we had to investigate alternative representations. For example CADP is based on a synchronous communication model (rendez-vous), whereas SDL communications are fully asynchronous (via queues).

Another motivation concerns time modeling. Finding a "reasonable" notion of time in asynchronous systems is a non trivial question and this is reflected by the variety of the existing proposals for existing FDTs. For instance, the SDL syntax defines a timer concept, but there is no consensus on how time progresses, and different SDL tools have adopted different choices. Similarly, in the original LOTOS definition there is no particular notion of time, whereas different timed extensions are currently being proposed [LL97,Que98]. Choosing an appropriate timed extension for an FDT should take into account not only technical considerations about the semantics of timed systems but also more pragmatic ones related to the appropriateness for use in a system engineering context. We believe that the different ideas about extensions of the language must be validated experimentally before being adopted to avoid phenomena of rejection by the users.

These problems motivated the development of IF, an intermediate representation for timed asynchronous systems. The requirements on such a formalism are the following ones:

- it must be sufficiently expressive to be used as an intermediate representation for the above mentioned specification formalisms, or at least for reasonably large subsets of them.
- it must have a formally defined operational semantics, and be flexible enough to experiment different choices and extensions.
- it must be supported by a set of well defined APIs, at different levels of program representation, allowing to interface existing validation tools and to experiment new ones.

The paper is organized as follows. First, we define the IF formalism, its main concepts and its operational semantics. We also discuss its expressiveness with respect to other models and specification formalisms, in particular regarding the timing aspects. Then, we present a set of tools interconnected within an open validation environment for IF specifications. Finally, we illustrate the use of IF on a small example, a distributed leader election algorithm on which different kinds of validation are performed.

2 Presentation of IF

In the following sections, we give a brief overview of the main features of IF, its operational semantics in terms of labeled transition systems. A more complete description of IF and of its semantics can be found in [BFG+98].

2.1 Syntax

An IF system is a set of processes communicating either asynchronously through a set of *buffers* or synchronously through a set of *gates*. The timed behavior of a system can be controlled through *clocks* (like in timed automata [ACD93,HNSY94]).

IF system definition: A system is a tuple *Sys* = (*glob-def*, PROCS, S) where

- *glob-def* = (*type-def*, *sig-def*, *gate-def*, *var-def*, *buf-def*) is a list of global definitions, where *type-def* is a list of type definitions (enumerated types, arrays, records and also abstract data types[1]) *sig-def* defines a list of parameterized *signals* (as in SDL), *gate-def* defines a list of parameterized *gates* (as in LOTOS), *var-def* is a list of global variables, and finally, *buf-def* is a list of *buffers* through which the processes communicate by asynchronous signal exchange (as in Promela[Hol91] or SDL). Notice that we allow various types of buffers: FIFO queues, stacks or bags, which can chosen to be unbounded or bounded and reliable or lossy.
- PROCS defines a set of processes described in the following paragraph.
- S is a synchronization expression, as in LOTOS or CSP, telling how the processes defined in PROCS synchronize. Such a synchronization expression is given by the following grammar where C is a (possible empty) set of gates:

  S ::= P∈PROCS | S [[C]] S

  Thus, a system S is either a process P or a parallel composition of two subsystems S_1 and S_2 with rendez-vous synchronization on the set of gates C. In a system of the form S_1 [[C]] S_2 transitions concerning a gate in C are executed synchronously in the two subsystems whereas all other transitions are interleaved.

IF process definition: Processes are defined by a set of local variables, a set of control states and a set of control transitions. A process P∈PROCS is a tuple P= (*var-def*, Q, CTRANS), where:

[1] where we suppose that the user provides also implementations of the introduced functions, otherwise expressions containing them are handled syntactically

- *var-def* is a set of local variable definitions including also clocks[2] (as in timed automata)
- `Q` is a set of **control states** on which the following attributes are defined:
  - *stable*(q) and *init*(q) are boolean attributes, where the attribute *stable* can be used to control the level of atomicity: only stable states are visible on the semantic level.
  - the attributes *save*(`q`), *discard*(`q`) are sets of `filters` of the form
    `signal-list` **in** `buf` **if** `cond`.
    which filter the buffers contents in this state. For example, *discard*(`q`) is used to eliminate silently unexpected signals: when consuming the next signal in the FIFO queue `buf`, all signals of `signal-list` preceding it are discarded in the same atomic step, if the boolean expression `cond` evaluates to `true`. These primitives are useful in practice and taken from SDL.
- cTRANS is a set of **control transitions**, between control states `q,q'`$\in$`Q`, which may be of the following types:
  - input transitions which are triggered by some signal read from one of the communication buffers (as in SDL) and internal transitions without input:

$$\mathtt{q} \xrightarrow[(\mathtt{urg})]{\mathtt{g} \mapsto \{\mathtt{input}\ ;\}\ \mathtt{body}} \mathtt{q}'$$

  - synchronization transitions which are executed simultaneously with compatible ones in other processes of the system (as in LOTOS):

$$\mathtt{q} \xrightarrow[(\mathtt{urg})]{\mathtt{g} \mapsto \mathtt{sync}} \mathtt{q}'$$

Where in all cases:

- `g` is a boolean *guard* of the transition which may depend on visible variables in the process (including clocks) and predefined tests on buffers content (e.g., emptiness).
- `urg`$\in$**{eager, delayable, lazy}** defines the *urgency type* of the transition. **eager** transitions have absolute priority over progress of time, **delayable** transitions may let time progress, but only as long as they remain enabled, whereas **lazy** transitions cannot prevent progress of time. These urgency types have been introduced in [BST98], which shows that the use of urgency predicates on transitions (instead of time progress conditions) facilitates the compositional specification of timed systems.
- `input` is an input of the form "**input** `sig`(*reference_list*) **from** `buf` **if** `cond`" where
  - `sig` is a signal,
  - *reference_list* the list of variables[3] (excluding clocks) in which the received parameters are stored,
  - `buf` is the name of the buffer from which the signal should be read

[2] one can also define timers (as in SDL) which can be set to any positive value, decrease with progress of time and expire if they reach the value *zero*; to simplify the presentation we do not include them in this document

[3] or "assignable" expressions such as elements of records or arrays

  - `cond` is a time independent "post guard" defining the condition under which the received signal is accepted and it usually depends on received parameters.

  Intuitively, an input transition is enabled if its guard is true, the first signal to be consumed (according to the attributes *save*(`q`) and *discard*(`q`)) is of the form `sig`$(v_1, ...v_k)$ and the post guard holds (after assigning the values $v_1, ...v_k$ to the variables of the *reference_list*)
- `sync` is a synchronization of the form "**sync** `gate` *comm_list* **if** `cond`" where
  - `gate` is a synchronization gate defined at system level,
  - *comm_list* is a list of *communication offers*:
    - either an output communication offer of the form `!exp`, where the expression `exp` represents the sent value
    - or a input communication offer of the form `?ref`, where `ref` is a local variable[3] in which the received value is stored.
  - `cond` is again a time independent post guard used to restrict the values that the process is willing to accept.

  The concept of synchronization is taken from LOTOS: the simultaneous execution of synchronization transitions concerning the same gate allows an instantaneous exchange of values between several processes. Notice that clock expressions cannot appear as communication offers.
- `body` is a sequence of atomic actions of the following types:
  - asynchronous *outputs* of the form "**output** `sig`(*par_list*) **to** `buf`" append a *signal* of the form "`sig`(*par_list*)" to the buffer `buf`.
  - usual *assignments* between discrete variables.
  - *settings* of clocks, which have the effect to assign to a clock a specific value.

2.2 Semantics

The semantics of IF is based on concepts taken respectively from LOTOS, SDL and timed automata. We define it by translating IF systems into timed automata with urgency [BST98]. First, we show how to associate a timed automaton with each process, and then, how two timed automata can be composed into a single one[4]. Such a timed automaton can then be interpreted either using discrete or dense time depending on the verification tools and properties considered. Notice that the discrete/dense interpretation of time does not influence the translation from IF to a timed automata.

Association of a Timed Automaton with a process: Let `P`$= ($*var-def*, `Q`, CTRANS$)$ be a process definition in the system `Sys` and furthermore:

- Let BUF be a set of buffer environments $\mathcal{B}$, representing possible contents of the buffers of the system, on which — depending on the declared buffer type — all necessary primitives are defined: e.g. "get the first signal of a given buffer, taking into account the *save* and the *discard* attributes of the control state", "append a signal at the end of a buffer", etc.

[4] Notice that the semantics is compositional in the sense that, in order to associate a timed automaton with a system one can also first compose the system into a unique process and then associate a timed automaton with this process

- Let ENV be a set of environments $\mathcal{E}$ defining the set of valuations of all discrete variables defined in the system `Sys` (the local and the global ones)

The semantics of the process `P` is the timed automaton [P] = (Q×ENV×BUF,TRANS) where

- Q×ENV×BUF is the set of states, for which we extend the attributes of control states in a natural manner, e.g. *tpc*(q,($\mathcal{E}$,$\mathcal{B}$)) is the partial evaluation of *tpc*(q) in ($\mathcal{E}$,$\mathcal{B}$). Notice that the set of data environments ENV can be split into $\text{ENV}_{loc} \times \text{ENV}_{glob}$ where ENV_{loc} concerns only local variables of the process and ENV_{glob} concerns the global variables of the system.
- TRANS is the set of transitions of the timed automaton obtained from control transitions by the following two rules:
  1. For any input or internal transition
     $$\texttt{q} \xrightarrow[(\texttt{urg})]{\texttt{g} \mapsto (\texttt{sig}(x_1...x_n), \texttt{buf}, \texttt{cond}) \ ; \ \texttt{body}} \texttt{q}' \in \text{cTRANS}$$
     and for any ($\mathcal{E}$,$\mathcal{B}$)∈VAL, the transition
     $$(\texttt{q}, (\mathcal{E}, \mathcal{B})) \xrightarrow[(\texttt{urg})]{\ell : \texttt{g}' \mapsto \texttt{body}'} (\texttt{q'}, (\mathcal{E}', \mathcal{B}')) \in \text{TRANS} \qquad \text{if}$$
     - g' is the the partial evaluation of g in ($\mathcal{E}$,$\mathcal{B}$), which is an expression depending only on clocks.
     - let $\mathcal{B}''$ be the buffer environment obtained after consuming $\texttt{sig}(v_1...v_n)$ in buffer `buf` (and after elimination of appropriate signals of the *discard*(q) attribute and saving of the signals of the *save*(q) attribute)
     - let $\mathcal{E}''=\mathcal{E}[v_1...v_n/x_1...x_n]$ is obtained by assigning v_i to x_i,
     - the post guard `cond` evaluates to `true` in the environment $(\mathcal{E}'', \mathcal{B}'')$
     - $(\mathcal{E}', \mathcal{B}')$ is obtained from $(\mathcal{E}'', \mathcal{B}'')$ by executing all the assignments of the `body`, and by appending all signals required by outputs in the `body`.
     - `body`' is the sequence of settings of clocks which remain as such in the timed automaton,
     - ℓ is an appropriate label used for tracing.
  2. For any synchronization transition of the form
     $$\texttt{q} \xrightarrow[(\texttt{urg})]{\texttt{g} \mapsto \texttt{c} : !\texttt{exp}_1 \ ?\texttt{y}_2 \ ... \ : \ \texttt{cond}} \texttt{q}' \in \text{cTRANS}$$
     and for any ($\mathcal{E}$,$\mathcal{B}$)∈VAL, the transition
     $$(\texttt{q}, (\mathcal{E}, \mathcal{B})) \xrightarrow[(\texttt{urg})]{\ell : \texttt{g}' \mapsto \text{skip}} (\texttt{q'}, (\mathcal{E}', \mathcal{B}')) \in \text{TRANS} \qquad \text{if}$$
     - g' is the the partial evaluation of g in ($\mathcal{E}$,$\mathcal{B}$),
     - the expression $\texttt{exp}_i$ evaluates to the value v_i in ($\mathcal{E}$,$\mathcal{B}$),
     - $\mathcal{E}'=\mathcal{E}[v_2...v_j.../y_2...y_j...]$ for some v_j belonging to the domain of y_j,
     - the post guard `cond` evaluates to `true` in the environment $\mathcal{E}$',
     - the label ℓ is equal to c $!\texttt{v}_1$ $!\texttt{v}_2$ ...

Composition of models: The timed automaton associated with a system of the form *Sys* = (*glob-def*,PROCS,S) is obtained by composing the timed automata of processes

according to the composition expression S. The composition rules presented correspond to the *and*-parallel composition described in [BST98].

Let $[\![P_i]\!] = (Q_i \times \text{VAL}, \text{TRANS}_i)_{i=1,2}$ be the timed automata associated with processes or subsystems of *Sys* — where VAL concerns only all global variables and is of the form $\text{ENV}_{glob} \times \text{BUF}$ and the valuations of local variables are integrated into the set of control states Q_i — and C a set of gates.
Then, $[\![P_1 [\![C]\!] P_2]\!] = [\![P_1]\!] \; [\![C]\!] \; [\![P_2]\!] = (Q \times \text{VAL}, \text{TRANS})$ where

- $Q = Q_1 \times Q_2$ where
$$\begin{aligned} \mathit{init}((q_1,q_2)) &= \mathit{init}(q_1) \wedge \mathit{init}(q_2) \\ \mathit{stable}((q_1,q_2)) &= \mathit{stable}(q_1) \wedge \mathit{stable}(q_2) \\ \mathit{tpc}((q_1,q_2)) &= \mathit{tpc}(q_1) \wedge \mathit{tpc}(q_2) \end{aligned}$$
- TRANS is the smallest set of transitions obtained by the following two rules: the first one applies to all transitions of TRANS_1 which are not synchronizations on gates in C and there is also a symmetric rule for transitions of TRANS_2.

$$\frac{(q_1, \mathcal{V}) \xrightarrow[(\mathtt{urg})]{\ell : \; g \mapsto \mathtt{body}} (q_1', \mathcal{V}') \in \text{TRANS}_1 \quad \text{and} \quad \neg \mathit{stable}(q_1) \vee \mathit{stable}(q_2)}{((q_1, q_2), \mathcal{V}) \xrightarrow[(\mathtt{urg})]{\ell : \; g \mapsto \mathtt{body}} ((q_1', q_2), \mathcal{V}') \in \text{TRANS}}$$

The requirement on stableness implies that there is no interleaving in non stable states; they are transient states, such that a finite sequence of transitions between to stable states can be considered as *one* atomic transition.
The second rule concerns the synchronizations on gates $c \in C$

$$\frac{\begin{array}{l}(q_1, \mathcal{V}) \xrightarrow[(\mathtt{u1})]{\ell : \; g_1 \mapsto \text{skip}} (q_1', \mathcal{V}') \in \text{TRANS}_1 \quad \text{and} \\ (q_2, \mathcal{V}) \xrightarrow[(\mathtt{u2})]{\ell : \; g_2 \mapsto \text{skip}} (q_2', \mathcal{V}') \in \text{TRANS}_2\end{array}}{((q_1, q_2), \mathcal{V}) \xrightarrow[(\mathtt{urg})]{\ell : \; g_1 \wedge g_2 \mapsto \text{skip}} ((q_1', q_2'), \mathcal{V}') \in \text{TRANS}}$$

In this rule, the synchronization of two transitions with the same urgency attribute result in a transition with this attribute, the composition of an **eager** transition with any other transition results in an **eager** one, an in order to compose a **lazy** with a **delayable** transition, one needs to decompose the delayable one into two transitions, an **eager** and a **lazy** one, which under a reasonable restriction is always possible [BST98].

The semantics of Timed Automata The model of time of IF is that of communicating timed automata with urgency introduced in [BST98]. Each process has a number of clocks which increase with progress of time (either in a discrete or continuous manner). Clocks can be "tested" in the guards and "set" in the bodies of the transitions. In this model, time is considered global, that is, it progresses synchronously in all processes of the system. The main problem is "when can time progress?". In timed automata

[ACD93], time progress is defined by means of "invariants" associated with each state, such that time is allowed to progress as long the invariant expression evaluates to true. The main problem with this model is that it allows not to express urgency of transitions. A model avoiding this problem is obtained by associating with every transition a *deadline* (a predicate stronger than the guard), meaning that, whenever the deadline predicate evaluates to true, the transition has priority over time progress. In [BST98], it has been shown that a much simpler model using just three possible *urgency* attributes, instead of deadlines, is sufficient: *eager* transitions have always priority over time, *delayable* transitions may let time progress, but only as long as they remain enabled, and *lazy* transitions cannot prevent time from progressing. This is the time model we have chosen in IF[5].

The semantics of timed automata with urgency is defined in [BST98]. Let $A = (\mathcal{Q}, \textsc{Trans})$ be a timed automaton. Let TIME be a set of environments for clocks, where $\mathcal{T} \in \textsc{time}$ defines for every clock a value in a time domain T (positive integers or reals). Setting a clock affects $\mathcal{T}$ by changing the value of the set clock to the specified value. Progress of time by an amount δ transforms the valuation $\mathcal{T}$ into the valuation $\mathcal{T} \boxplus \delta$ in which the values of all clocks are increased by δ.

The semantics of A is defined by the labeled transition system $(\mathcal{Q} \times \textsc{time}, \rightarrow)$ where the transition relation $\rightarrow$ consists of two types of transitions, discrete ones and time progress transitions:

- For any transition $\mathtt{q}_1 \xrightarrow[(\mathtt{urg})]{\ell : \mathtt{g} \mapsto \mathtt{body}} \mathtt{q}_2 \in \textsc{Trans}$ and $\mathcal{T} \in \textsc{time}$, there exists a discrete transition of the form $(\mathtt{q}_1, \mathcal{T}) \xrightarrow{\ell} (\mathtt{q}_2, \mathcal{T}')$ if
  - the guard g evaluates to `true` in $\mathcal{T}$,
  - and $\mathcal{T}$' is obtained from $\mathcal{T}$ by executing all clock settings of the `body`.
- in any state ($\mathtt{q}$,$\mathcal{T}$), time can progress by the amount δ, that is
  $(\mathtt{q}, \mathcal{T}) \xrightarrow{\text{time: } \delta} (\mathtt{q}, \mathcal{T} \boxplus \delta)$
  if time can progress in all states ($\mathtt{q}$,$\mathcal{T} \boxplus \delta'$) for $0 \leq \delta' < \delta$, where time can progress in a state ($\mathtt{q}$,$\mathcal{T}$) if and only if the following conditions hold:
  - *stable*($\mathtt{q}$), that means time progresses only in stable, never in transient states
  - *no* **eager** transition is enabled in ($\mathtt{q}$,$\mathcal{T}$)
  - for each **delayable** transition `tr` enabled in ($\mathtt{q}$,$\mathcal{T}$), there exists a positive amount of time ϵ, such that `tr` remains enabled when time progresses by ϵ. That means enabled delayable transitions allow time to progress, but only as long as they remain enabled.

3 IF and other formalisms

IF is a formalism for the description of asynchronous systems at a programming language level. It has not been designed with the aim to replace specification languages

[5] in order to include the model of timed automata, we allow associate an explicit *tpc* attribute with control states, but we don't present this feature here

such as LOTOS, SDL, but as a general intermediate representation for them, in particular for SDL. The expressiveness of IF and its adaptedness as an intermediate representation for some specification formalisms are discussed below.

3.1 SDL

The definition of SDL (Specification and Description Language) started in 1974 and it has been standardized by CCITT in 1988 [IT94]. SDL is based on extended finite state machines communicating asynchronously via queues. There exists a formal semantics of the language defined in [IT94], various authors have criticized it and proposed alternative ones, such as [Bro91,BMU98,God91] to name only a few of them. Currently, SDL is widely accepted by the industrial community. This is due mainly to the fact that SDL development is supported by methodologies [OFMP+94] and commercial tools [Ver96,AB93] in all phases from requirement analysis, design and validation to implementation. However, the validation capabilities of the commercial tools are rather limited with respect to the ones existing in the academic community.

There is no standard semantics of time defined for SDL and each tool uses its own. For example, *Object*GEODE uses a very "synchronous" time concept where time can only progress when the system is blocked (in terms of IF that means all transitions are eager), whereas in other tools time can always progress (all transitions are lazy). This shows that the currently implemented notions of time of SDL are extreme ones — which is often considered as problem by the users, and leads to unnecessary complicated descriptions — and many intermediate solutions are possible using IF as discussed in previous section.

We have identified a large static subset of SDL which we are able to translate into IF. That is, with the exception of dynamic creation of processes and some mobility aspects of communication, we can define a syntactic level translation between these two formalisms. A prototype translator has been implemented using the SDL/API Interface provided by ObjectGeode [Ver96]. More detailed information about it can be found in [BFG+99b].

3.2 LOTOS

LOTOS (Language Of Temporal Ordering Specifications) [BB88] has been developed and standardized by ISO in 1989. It is a process-algebra based on CCS[Mil80] and CSP[Hoa84]. In LOTOS, the communication is synchronous using rendez-vous. LOTOS has a well-defined operational semantics and there exist several tools supporting it.

The right approach to model and validate LOTOS specifications is recognized to be the use of Petri nets, rather than communicating extended automata [GS90] as intermediate representation. However, our experience with LOTOS has shown that often the specifications have the form of a parallel composition of sequential components (processes). This observation motivated also the use of compositional generation methods, which gives good results for this kind of LOTOS specifications [KM97].

The timed extensions introduced in E-Lotos[Que98], ET-Lotos[LL97] and Lotos-NT[Sig99] are similar to that of IF, only that the urgency of an action is defined implicitly by its type: "exceptions" and internal actions are urgent, whereas observable

actions are not. This is due to the fact, that they aim for a much stronger form of compositionality, where with each process can be associated directly a labeled transition system (and not a timed automaton) which then can be composed to a system model.

We plan to investigate the translation of decomposable LOTOS specifications into IF, as parallel composition with synchronization between processes can be handled in IF. Furthermore, a reasonably small Petri Net (corresponding to a non-decomposable LOTOS part) can be modeled as an IF process.

3.3 PROMELA

Another language we have considered is Promela, the native language of the Spin model-checker [Hol91]. It has not been designed as a specification language but as an intermediate representation language for protocols, mainly for validation purposes. It is based on extended finite-state machines communicating asynchronously or synchronously via queues. We consider Promela as it has a relatively important visibility as well in the academic community as in the industrial one. Its success is due to the high availability of Spin, which provides powerful model-checking algorithms based on partial-order reductions.

There exist timed extensions of Promela. The one proposed in [CT96] is based on timed automata, whereas the one of [BD98] has a similar time concept as *Object*GEODE: all set timers decrease synchronously until one of them expires; then time is blocked until the corresponding timeouts are consumed, where these timeout consumptions take place when no other transition is possible in the system.

A translator from IF to Promela has been developed in the framework of the VIRES Esprit-LTR project at Eindhoven University and has been used to verify SDL specification with Spin [BDHS99]. We plan to study also the translation from Promela to IF. As for SDL, there are some limitations due to dynamic process creation and mobility features of Promela.

4 A validation environment based on IF

One of the main motivations for developing the IF intermediate representation is to provide an "open" validation environment, able to make heterogeneous tools cooperate within a single framework. Especially for SDL, solid industrial tools for editing and code generation have been built which are used by a large community of users. On the other hand, there exist many verification tools built upon diverse formalisms — such as the Spin tool [Hol91] based on Promela, the CADP tool[FGK+96] based on LOTOS, the SMV tool [McM93] based on extended automata, tools for the verification of timed systems such as KRONOS[Yov97] and Uppaal [LPY97] based on different representations of timed automata, to name only a few of them.

An integrated validation environment should fulfill the following requirements:

- First of all, it is able to support several validation techniques, from symbolic interactive simulation to automatic property checking, together with test case generation and executable code generation. Indeed, all these functionalities cannot be embodied in a single tool and only tool integration facilities can provide all of them.

- Moreover, for a sake of efficiency, this environment also has to support several level of representations. For instance it is well-known that model-checking verification of real life case studies usually needs to combine several optimization techniques to overcome the state explosion problem.

  In particular, some of these techniques rely on a syntactic level representation, like static analysis and computations of abstractions (for which it may be necessary to cooperate with decision procedures or a theorem-prover). Other techniques operate on a representation of the underlying semantic model, such as on-the-fly analysis, bisimulation based model reduction or model-checking. These representations can be either implicit, enumerative or symbolic and are explained below.
- Another important feature is to keep this environment *open* and *evolutive*. Therefore, tool connections are performed by sharing either input/output file formats, or libraries of internal data structures. For this purpose several well-defined interfaces (APIs) must be provided.

In the remainder of the section we present the overall architecture of the already existing environment and some of its related components. In the future new connections with existing tools and new analysis modules may be added. Figure 4 describes the existing (plain arrows) and planned (dashed arrows) connections of this environment. However this figure does not represent which tools are part of the IF distribution and which are interconnections with other existing tools.

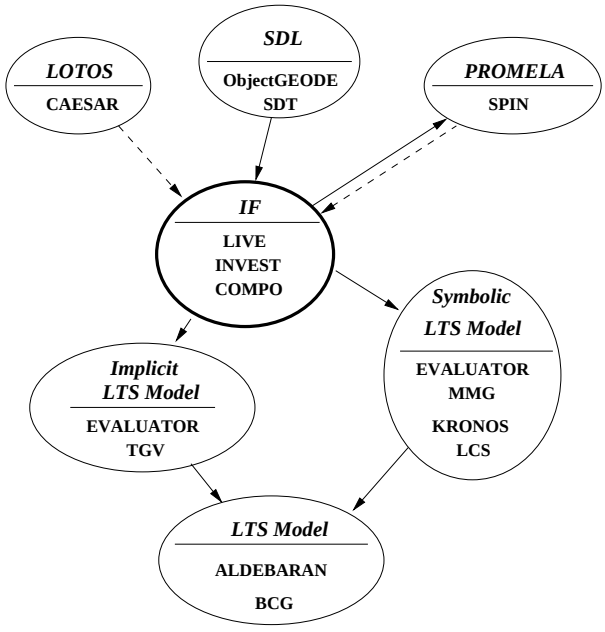

Fig. 1. An open validation environment for IF

4.1 Overall architecture

Our IF validation environment is built upon two levels of program representations, each of them being accessed through a well-defined API.

The syntactic level allows to consult and modify the abstract tree on an IF program. Since all the variables, timers, buffers and the communication structure are still explicit, high-level transformations based on static analysis (such as *live variable* computation) or program abstraction can be applied. Moreover, this API is also well suited to implement translators between IF and other specification formalisms.

The execution model level gives access to the LTS representing the semantics of the IF program. The following three APIs are those offered in CADP for different types of representations. In the IF environment, also mixed representations are used.

- The **implicit enumerative representation** is based on the OPEN-CAESAR [Gar98] philosophy. It consists in a set of C functions and data structures allowing to compute on demand the successors of a given state. This piece of C code is generated by the IF compiler, and it can be linked with a "generic" exploration program performing on the fly analysis (deadlock detection, model-checking, test-case generation, ...).
- In the **symbolic representation** (called SMI [Boz97]) sets of states and transitions of the LTS are expressed by their characteristic functions over a set of finite variables. These functions are implemented in terms of decision diagrams (BDDs [Bry86] and MDDs). Existing applications based on this API are symbolic model-checking and minimal model generation.
- Finally, the **explicit enumerative representation** simply consists in an LTS file format with an associated access library. Although such an explicit representation is not suitable for handling large systems globally, it is still useful in practice to minimize some of its abstractions with respect to bisimulation based relations (like in *compositional generation*, see below).

Below, we discuss the tools being part of the IF verification environment and some external tools for which exists a strong connection.

CADP [FGK+96,BFKM97] is a tool set for the verification of LOTOS specifications. It has been developed and by VERIMAG and the VASY team of INRIA Rhône-Alpes. We briefly present here two of its verifiers which are also part of the IF environment:

- ALDEBARAN compares and minimizes finite LTSs with respect to various *simulation* or *bisimulation* relations. This allows the comparison of the observable behavior of a specification with its expected one, described at a more abstract level.
- EVALUATOR is a "on-the-fly" model-checker for formulas of the alternating-free μ-calculus [Koz83].

MMG [FKM93], developed at VERIMAG is a minimization tool based on a partition refinement algorithm combined with a reachable state space computation [BFH90]. This tools works on the symbolic SMI interface.

ObjectGEODE [Ver96] is a commercial tool set developed by VERILOG supporting SDL, MSC and OMT. It includes graphical editors and compilers for each of these formalisms. It also provides a C code generator and a simulator to help the user to interactively debug an SDL specification.

*Object*GEODE also provides an API offering a set of functions and data structures to access the abstract tree generated from an SDL specification. Our translation tool (SDL2IF) uses this abstract tree to generate an operationally equivalent IF specification.

KRONOS [Yov97], developed at VERIMAG is a model-checker for symbolic verification of TCTL formulae on communicating timed automata. The current connection with the IF/CADP environment is as follows: control states and discrete variables are expressed using the IF/CADP implicit enumerative representation whereas clocks are expressed using an appropriate symbolic representation (particular polyhedra). Currently we are working on a more efficient translation of SDL timers into clocks.

TGV [FJJV97] is a test sequence generator built upon CADP jointly by VERIMAG and the PAMPA project of IRISA. TGV aims to automatically generate test cases for conformance testing of distributed systems. Test cases are computed during the exploration of the model and they are selected by means of *test purposes*. Test purposes characterize some abstract properties that the system should have and one wants to test, given trees of labels, decorated with verdicts "ok" and "fail".

INVEST [BLO98] is a symbolic verification tool based on the interaction with the theorem prover PVS [OSR93] computing abstractions and invariants on a set of guarded command processes communicating through shared variables. It has been developed jointly by VERIMAG, the university of Kiel and SRI. We have implemented translations between this formalism and IF, allows us to compute abstract systems.

LIVE [BFG99a] is a tool developed at VERIMAG. It transforms an IF specifications into an equivalent IF specification with a smaller state graph by means of static analysis. Presently, only simple algorithms, such as constant variable elimination and dead variable resetting (a variable which at some control point is never used before assigned again, is set to some default value) are implemented. Even this very simple analysis is very efficient, as a reduction of the state space by a factor 100 is common. In the future, we intend also to implement algorithms building weaker abstractions, for example elimination of irrelevant variables.

COMPO is a tool being developed at VERIMAG for compositional generation of minimal models associated with IF programs. This compositional generation method has already been applied for specification formalisms based on *rendez-vous* communication, and has been shown efficient in practice [GLS96,Val96,KM97]. It has not been investigated for systems based on communication via buffers, may be, because buffers raise several difficulties or due to the lack of suitable representations and tools. The potential benefit of this approach is illustrated on an example in the next section.

5 An illustrating example

We present a simple example to illustrate the IF formalism and related verification tools. We consider a *token ring*, that is a system of n stations (processes) $S_1, \ldots S_n$, connected in a circular network, in which a station is allowed to access some shared resource R only when it "owns" a particular message, the *token*. If the network is unreliable it is necessary to recover from token loss. This can be done using a *leader election algorithm* [Lan77,CR79] to determine a station responsible for generating a new token.

```
signal
    close(pid);
    open(pid);
    claim(pid, bool);
    token;
buffer
    Q1 : queue :lossy of claim, token;
    Q2 : queue :lossy of claim, token;
    Q3 : queue :lossy of claim, token;
    Q4 : queue :lossy of claim, token;
sync
    S1 ||| S2 ||| S3 ||| S4
end;
process S1;
var
    worried : timer;
    round, rnd: bool;
    adr: pid;
...
```

Table 1. IF global definitions

Table 1 shows the global definitions of the IF specification corresponding to the particular protocol considered in [GM97]. The signals `open` and `close` denote the access and the release of the shared resource (here a part of the environment). The signals `token` and `claim` are the messages circulating on the ring.

All stations S_i are identical up to their identity and described by an IF process as the one of Figure 2. The timer `worried` is set when the station waits for the token and reset when it receives it. On expiration of the timer `worried` token loss is assumed and an election phase is started. The "alternating bit" `round` is used to distinguish between valid claims (emitted during the current election phase) and old ones (cancelled by a token reception). In the `idle` state, a station may either receive the token from its neighbor (then it reaches the `critical` state and can access the resource) or receive the timer expiration signal (then it emits a claim stamped with its `address` and the current value of `round`) or receive a claim from its neighbor. A received claim is "filtered" if its associated `address` is smaller than its own address and transmitted unchanged if it is greater. If its own valid claim is received, this station becomes elected and generates a new token.

Model generation: We summarize in Table 2 the size of the models obtained from the token-ring protocol using three generation methods: directly from the initial IF program (global generation), using the live variable reduction (global + live) and using a compositional generation strategy (compositional + live).

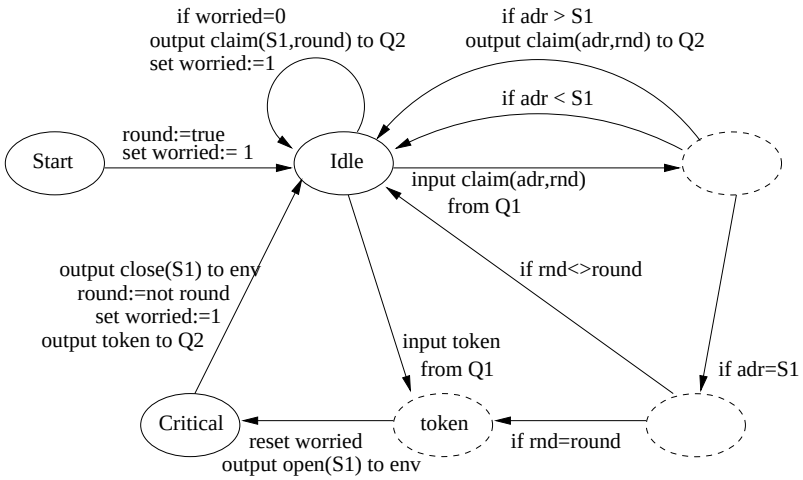

Fig. 2. The behavior of station S_1

The most spectacular reduction is obtained by the live reduction: the reduced model is about 100 times smaller than the one obtained by simultaneous generation, while preserving *all* properties (models 1 and 2 are strongly bisimilar). This is explained by the fact that only a few variables are live in each state: in the `idle` state the live variables are `round` and `worried`, in the `critical` state only `round` is live, while variables `adr` and `rnd` are never live.

model	*generation method*	*states*	*transitions*
1.	global	537891	2298348
2.	global + live	4943	19664
3.	compositional + live	1184	4788

Table 2. Models obtained for the token ring example

More reduction is achieved by the following compositional generation strategy yielding an LTS branching bisimilar to the original one:

1. We split the IF description into two parts, the first one contains processes S_1 and S_2 and the second one processes S_3 and S_4. For each one of these descriptions, the internal buffer between the two processes is *a priori* bounded to two places. Note that, when a bounded buffer overflows during simulation, a special *overflow* transition occurs in the corresponding execution sequence.

2. The LTS associated with each description is generated considering the "most general" environment providing any potential input. As `claim` and `token` can be transmitted at any time, *overflow* transitions appear in the generated LTSs.
3. In each LTS the input and output transitions relative to the internal buffers (Q_2 and Q_4) are hidden (i.e., renamed to the special τ action); then the two LTSs are reduced w.r.t an equivalence relation preserving the properties under verification. For the sake of efficiency we have chosen the branching bisimulation [vGW89] preserving all the safety properties (e.g. mutual exclusion).
4. The reduced LTSs are then translated back into an IF process (without variables), and the resulting processes are combined into a single global IF description with only two buffers (Q_1 and Q_3). It turns out that the LTS generated from this new description contains no *overflow* transitions (they have been cut off during the second composition, which confirms the hypothesis on the maximal size of the internal buffers).

Verification: We are interested in checking that the shared resource is accessed in mutual exclusion. For this, we consider as visible only the `open` and `close` actions.

Mutual exclusion property can be rephrased as follows: *after every* $\texttt{open}(S_i)$ *(station* i *enters the critical section) the only possible visible action is* $\texttt{close}(S_i)$ *(station* i *leaves the critical section) possibly after a number of internal moves* τ. This property can be expressed in the μ-calculus (see below) and verified with EVALUATOR, on any of the generated models.

$$\bigwedge_{i=S_1}^{S_4} \nu X.\ ([\texttt{open}(S_i)]\ \neg\mu Y.\ (\langle\overline{\{\texttt{close}(S_i), \tau\}}\rangle T\ \vee\ \langle\tau\rangle Y)\ \wedge\ [*]X)$$

Another approach to verify mutual exclusion is to compare the model of the specification with an abstract one expressing the desired behavior. For instance, after hiding of all actions different from $\texttt{close}(S_i)$ and $\texttt{open}(S_i)$, the minimal model for branching bisimulation of our example specification is the one shown in Figure 3. The reductions and comparisons have been carried out using ALDEBARAN.

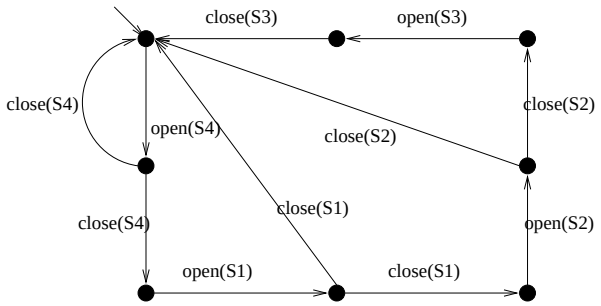

Fig. 3. The reduced behavior of the token ring.

Test Generation: We illustrate the use of TGV to extract test cases for the token ring protocol. We want to test the property stating that a station filters a received claim with a smaller address than its own and transmits it unchanged if it is greater. We chose a test purpose expressing that after S_4 has sent its claim, it will be transmitted unchanged by station S_1, then by S_2 and finally by S_3. The generated test case is shown in figure 4.

Test Case Dynamic Behaviour					
Test Case Name : castest					
Group :					
Purpose :					
Default :					
Comments :					
Nr	Label	Behaviour Description	Cts Ref	Verdict	C
1		s3? claim	claim3	INCONC	
2		s2? claim	claim2	INCONC	
3		s1? claim	claim1	INCONC	
4		s4? claim	claim0		
5		s1! claim, St tclaim	claim0		
6		s3? claim, Cl tclaim	claim3	INCONC	
7		s2? claim, Cl tclaim	claim2	INCONC	
8		s1? claim, Cl tclaim	claim1	INCONC	
9		s1? claim, Cl tclaim	claim4		
10		s2! claim, St tclaim	claim4		
11		s3? claim, Cl tclaim	claim3	INCONC	
12		s2? claim, Cl tclaim	claim2	INCONC	
13		s1? claim, Cl tclaim	claim1	INCONC	
14		s2? claim, Cl tclaim	claim5		
15		s3! claim, St tclaim	claim5		
16		s3? claim, Cl tclaim	claim3	INCONC	
17		s2? claim, Cl tclaim	claim2	INCONC	
18		s1? claim, Cl tclaim	claim1	INCONC	
19		s3? claim, Cl tclaim	claim6	(PASS)	
20		? tclaim		FAIL	
21		? tclaim		FAIL	
22		? tclaim		FAIL	

Fig. 4. TTCN test case

6 Conclusion and perspectives

We have presented the formalism IF which has been designed as an intermediate representation for SDL, but it can be used as a target language for other FDT as it contains most of the concepts used in these formalisms. The use of IF offers several advantages:

- IF has a formal semantics based on the framework of communicating timed automata. It has powerful concepts interesting for specification purposes, such as different urgency types of transitions, synchronous communication, asynchronous communication through various buffer types (bounded, unbounded, lossy, ...), and communications through shared variables.
- IF programs can be accessed at different levels through a set of well defined APIs. These include not only several low-level model representations (symbolic, enumerative, ...) but also higher level program representation, where data and communication structures are still explicit. Using these APIs several tools have been already interconnected within an open environment able to cover a wide spectrum of validation methods.

Our translator from SDL to IF has already been used successfully to analyze real-life SDL specifications with CADP and SPIN, and is actually being used to experiment different semantics of time for SDL using the connection with the KRONOS tool.

A concept which is not provided in IF is dynamic creation of new process instances of processes and parameterization of processes; this is due to the fact that in the framework of algorithmic verification, we consider only static configurations. However, it is foreseen in the future to handle some kind of parameterized specifications and to translate also systems with bounded process creation.

The results obtained using the currently implemented static analysis and abstractions methods are very encouraging. For each type of analysis, we built a module taking an IF specification as input and which generates a *reduced* one. This architecture allows to chain several modules to benefit in a modular way from multiple reductions applied to the same initial specification. We envisage to experiment more sophisticated analysis, such as constraints propagation, and more general abstraction techniques. This will be achieved either by developing dedicated components or through the connections with INVEST.

The IF package is available at `http://www-verimag.imag.fr/DIST_SYS/IF`.

References

[AB93] Telelogic AB. *SDT Reference Manual.* http://www.telelogic.se/solution/tools/sdt.asp, 1993.

[ACD93] R. Alur, C. Courcoubetis, and D.L. Dill. Model Checking in Dense Real Time. *Information and Computation*, 104(1), 1993.

[BB88] T. Bolognesi and E. Brinksma. Introduction to the ISO Specification Language LOTOS. *ISDN*, 14(1), jan 1988.

[BD98] D. Bošnački and D. Dams. Integrating Real Time into Spin: A Prototype Implementation. In *Proceedings of the FORTE/PSTV XVIII Conference*, 1998.

[BDHS99] D. Bošnački, D. Dams, L. Holenderski, and N. Sidorova. Verifying the MASCARA Protocol in SPIN. submitted to the SPIN'99 Workshop, mai 1999.

[BFG+98] M. Bozga, J-C. Fernandez, L. Ghirvu, S. Graf, L. Mounier, J.P. Krimm, and J. Sifakis. The Intermediate Representation IF. Technical report, Verimag, 1998.

[BFG99a] M. Bozga, J.-C. Fernandez, and L. Ghirvu. State Space Reduction based on Live Variables Analysis. In *Proceedings of SAS'99, Venezia, Italy*, LNCS, September 1999. to appear.

[BFG+99b] M. Bozga, J.-C. Fernandez, L. Ghirvu, S. Graf, J.P. Krimm, L. Mounier, and J. Sifakis. IF: An Intermediate Representation for SDL and its Applications. In *Proceedings of SDL-FORUM'99, Montreal, Canada*, June 1999.

[BFH90] A. Bouajjani, J.-C. Fernandez, and N. Halbwachs. Minimal Model Generation. In *Proceedings of CAV'90, Rutgers, New Jersey*, volume 3 of *DIMACS*, pages 85–92, June 1990.

[BFKM97] M. Bozga, J.-C. Fernandez, A. Kerbrat, and L. Mounier. Protocol Verification with the Aldebaran Toolset. *STTT*, 1(1+2):166–183, December 1997.

[BLO98] S. Bensalem, Y. Lakhnech, and S. Owre. Computing Abstractions of Infinite State Systems Compositionally and Automatically. In *Proceedings of CAV'98, Vancouver, Canada*, volume 1427 of *LNCS*, June 1998.

[BMU98] J.A. Bergstra, C.A. Middelburg, and Y.S. Usenko. Discrete Time Process Algebra and the Semantics of SDL. Technical Report SEN-R9809, CWI, June 1998.

[Boz97] M. Bozga. SMI: An Open Toolbox for Symbolic Protocol Verification. Technical Report 97-10, Verimag, Sep 1997.

[Bro91] M. Broy. Towards a Formal Foundation of the Specification and Description Language SDL. *Formal Aspects on Computing*, 1991.

[Bry86] R.E. Bryant. Graph Based Algorithms for Boolean Function Manipulation. *IEEE Transactions on Computation*, 35(8), 1986.

[BST98] S. Bornot, J. Sifakis, and S. Tripakis. Modeling Urgency in Timed Systems. In *International Symposium: Compositionality - The Significant Difference, Holstein, Germany*, volume 1536 of *LNCS*, 1998.

[CR79] E. Chang and R. Roberts. An Improved Algorithm for Decentralized Extrema-Finding in Circular Configurations of Processes. *Communications of ACM*, 22(5), 1979.

[CT96] C. Courcoubetis and S. Tripakis. Extending Promela and Spin for Real Time. In *Proceedings of TACAS'96*, volume 1055 of *LNCS*, 1996.

[FGK+96] J.-C. Fernandez, H. Garavel, A. Kerbrat, R. Mateescu, L. Mounier, and M. Sighireanu. CADP: A Protocol Validation and Verification Toolbox. In *Proceedings of CAV'96, New Brunswick, USA*, volume 1102 of *LNCS*, July 1996.

[FJJV97] J.-C. Fernandez, C. Jard, T. Jéron, and C. Viho. An Experiment in Automatic Generation of Test Suites for Protocols with Verification Technology. *SCP*, 29, 1997.

[FKM93] J.-C. Fernandez, A. Kerbrat, and L. Mounier. Symbolic Equivalence Checking. In *Proceedings of CAV'93, Heraklion, Greece*, volume 697 of *LNCS*, 1993.

[Gar98] H. Garavel. OPEN/CÆSAR: An Open Software Architecture for Verification, Simulation, and Testing. In *Proceedings of TACAS'98, Lisbon, Portugal*, volume 1384 of *LNCS*, 1998.

[GLS96] S. Graf, G. Lüttgen, and B. Steffen. Compositional Minimisation of Finite State Systems using Interface Specifications. *Formal Aspects of Computation*, 3, 1996.

[GM97] H. Garavel and L. Mounier. Specification and Verification of Distributed Leader Election Algorithms for Unidirectional Ring Networks. *SCP*, 29, 1997.

[God91] J.C. Godskesen. An Operational Semantic Model for Basic SDL. Technical Report TFL RR 1991-2, Tele Danmark Research, 1991.

[GS90] H. Garavel and J. Sifakis. Compilation and Verification of LOTOS Specifications. In *Proceedings of the 10th PSTV, Ottawa, Canada*, June 1990.

[HNSY94] T.A. Henzinger, X. Nicollin, J. Sifakis, and S. Yovine. Symbolic Model Checking for Real-Time Systems. *Information and Computation*, 111(2), 1994.

[Hoa84] C.A.R. Hoare. *Communicating Sequential Processes.* Prentice Hall International, 1984.

[Hol91] Gerard J. Holzmann. *Design and Validation of Computer Protocols.* Prentice Hall Software Series, 1991.

[ISO88] ISO/IEC. LOTOS — A Formal Description Technique Based on the Temporal Ordering of Observational Behaviour. Technical Report 8807, International Organization for Standardization — Information Processing Systems — Open Systems Interconnection, 1988.

[IT94] ITU-T. *Recommendation Z-100. Specification and Description Language (SDL) and Annexes F.2: Static Semantics and F.3: Dynamic Semantics.* 1994.

[KM97] J.P. Krimm and L. Mounier. Compositional State Space Generation from LOTOS Programs. In *Proceedings of TACAS'97, Enschede, The Netherlands*, volume 1217 of *LNCS*, 1997.

[Koz83] D. Kozen. Results on the Propositional μ-Calculus. In *Theoretical Computer Science*. North-Holland, 1983.

[Lan77] G. Le Lann. Distributed Systems – Towards a Formal Approach. In *Information Processing 77*. IFIP, North Holland, 1977.

[LL97] L. Leonard and G. Leduc. An Introduction to ET-LOTOS for the Description of Time-Sensitive Systems. *Computer Networks and ISDN Systems*, (29), 1997.

[LPY97] K.G. Larsen, P. Petterson, and W. Yi. UPPAAL: Status & Developments. In *Proceedings of CAV'97, Haifa, Israel*, volume 1254 of *LNCS*, 1997.

[McM93] K.L. McMillan. *Symbolic Model Checking: an Approach to the State Explosion Problem*. Kluwer Academic Publisher, 1993.

[Mil80] R. Milner. A Calculus of Communication Systems. In *LNCS*, number 92. 1980.

[OFMP+94] A. Olsen, O. Færgemand, B. Møller-Pederson, R. Reed, and J.R.W. Smith. *Systems Engineering Using SDL-92*. North-Holland, 1994. ISBN 0444 898727.

[OSR93] S. Owre, N. Shankar, and J.M. Rushby. A Tutorial on Specification and Verification Using PVS. Technical report, Computer Science Laboratory, SRI International, February 1993.

[Que98] J. Quemada. Final Comitee Draft on Enhancements to LOTOS. Technical report, ISO/IEC JTC1/SC33/WG9, April 1998.

[Sig99] M. Sighireanu. *Contribution at the Definition and Implementation of E-LOTOS*. PhD thesis, Université Joseph Fourier, Grenoble, 1999.

[Val96] A. Valmari. Compositionality in State Space Verification. In *Application and Theory of Petri Nets*, volume 1091 of *LNCS*, 1996.

[Ver96] Verilog. Object*GEODE SDL Simulator - Reference Manual*. http://www.verilogusa.com/solution/pages/ogeode.htm, 1996.

[vGW89] R.J. van Glabbeek and W.P. Weijland. Branching-Time and Abstraction in Bisimulation Semantics. CS R8911, CWI, 1989.

[Yov97] S. Yovine. KRONOS: A Verification Tool for Real-Time Systems. *STTT*, 1(1-2), Dec 1997.