Hierarchical states in the Compositional Interchange Format

H. Beohar  D.E. Nadales Agut  D.A. van Beek  P.J.L. Cuijpers
Eindhoven University of Technology (TU/e)
P.O. Box 513, NL-5600 MB Eindhoven, The Netherlands
{H.Beohar, D.E.Nadales.Agut, D.A.van Beek, P.J.L. Cuijpers}@tue.nl

CIF is a language designed for two purposes, namely as a specification language for hybrid systems and as an interchange format for allowing model transformations between other languages for hybrid systems. To facilitate the top-down development of a hybrid system and also to be able to express models more succinctly in the CIF formalism, we need a mechanism for stepwise refinement. In this paper, we add the notion of hierarchy to a subset of the CIF language, which we call hCIF. The semantic domain of the CIF formalism is a hybrid transition system, constructed using structural operational semantics. The goal of this paper is to present a semantics for hierarchy in such a way that only the SOS rules for atomic entities in hCIF are redesigned in comparison to CIF. Furthermore, to be able to reuse existing tools like simulators of the CIF language, a procedure to eliminate hierarchy from an automaton is given.

1 Introduction

Modeling languages for hybrid systems, and hybrid automata in particular, are designed to combine computational and physical aspects of a system in one formal model. The compositional interchange format (CIF), presented in [3][14], is a hybrid modeling language based on hybrid automata [9], but with the semantics defined via structural operational semantics (SOS) [12] rules. One of the primary aims of CIF is to establish inter-operability among a wide range of tools by means of model transformations to and from the CIF. In addition, it is possible to specify hybrid systems using CIF, and perform simulation. The reason for using an SOS semantics in an automaton-based framework is that the model transformations to and from CIF are not only to be executed on ‘complete’ models, but also on components of bigger models. Thus, it is crucial that bisimulation (equivalence) is a congruence for all the constructs of the CIF. This is guaranteed using the process-tyft format of [11].

The CIF language contains the following features.

- Predicates in the locations of the automata that constrain the initial values and/or initial locations (init predicate), time behavior (invariants and time can progress predicates), and action behavior (invariants).
- Communication among automata using channels and shared variables.
- Scoping operators for declaring variables, actions, and channels.
- An initialization operator for restricting the initial conditions of variables. This allows initialization on a more global level as compared to the init predicates of automata.
- A synchronization operator for executing actions synchronously in parallel automata.
- An urgency operator for declaring actions or channels as urgent.

To facilitate the top-down development of a hybrid system in the CIF formalism we need a mechanism for stepwise refinement. In this work we develop such hierarchical extensions for a subset of CIF called
hCIF\textsuperscript{C}, in which we leave out the constructs for scoping, initialization, synchronization and urgency. Nevertheless, the semantics presented here is general enough to allow us to incorporate these concepts in a straightforward manner. In a later phase, we plan to extend hCIF\textsuperscript{C} to contain these constructs again. This requires us to already take into account some semantic features that are particular for the CIF (like the so-called variable trajectories, guard trajectories and termination trajectories that we discuss further on in this paper) while other features only appear in a reduced form (like the so-called environment transitions that are only used for establishing termination, and not for establishing consistency of a state).

Stepwise refinement is a framework for designing a system correct by construction. The following steps are involved in the stepwise refinement framework as laid out in [1]. One starts with design of a system at a higher level of abstraction, and usually the model designed at this level is called an abstract model. Then a concrete model is designed by adding more behavior into the abstract model such that the concrete model is a refinement of the abstract model. This process of refining is performed until a desired implementation is reached. Thus, any formalism that supports stepwise refinement, must incorporate the following two main things. First, it should provide a way to add new details in an abstract model. Secondly, it should provide \textit{at-least} sufficient conditions under which a concrete model is a refinement of a given abstract model by construction. In this paper, we consider only the first aspect of stepwise refinement.

Consequently, we introduce a notion of hierarchy in hCIF\textsuperscript{C} that allows a straightforward way to add new behavior to a given model. In the past, the following techniques were proposed in order to accommodate stepwise refinement in other formalisms.

- **Action refinement** [19]. In this approach an action in the alphabet of a process or an automaton is substituted by another process/automaton. However, the setting of action refinement is incompatible with the interleaving models of concurrency as pointed out in [15]. Since the CIF and hCIF\textsuperscript{C} formalisms are based on interleaving models of concurrency, we disregard this technique of refinement.

- **Statecharts** [8]. Statecharts were the first formalism that extended finite state machines with the concept of hierarchy. Conventionally, the semantics of statecharts requires a tree-structure on the set of locations of a statechart. Consequently, additional concepts from tree-structures, like least common ancestors, children of a location, etc., make the semantics complicated. We show in the current work that these additional concepts are unnecessary when reverting to a structural operational semantics. We only need to introduce the notion of a substructure. Other concepts of state-structures, like AND-states and OR-states, can be expressed through the \textit{parallel composition} and the \textit{multiple initial locations} of substructures, respectively. The concepts of history retention and inter-level transitions (not considered in this paper) are not supported directly in hCIF\textsuperscript{C}, but they can be emulated.

- **Hierarchical timed automata** [6, Chap 4.] are the extensions of statecharts with a finite set of clock variables modeling real time. Again the semantics of this formalism is based on the concepts of tree structures and for this reason we also disregard this approach. However, there is a common intuition about the passage of time in [6] with the current work. The time can pass in a hierarchical structure only if the time can pass in all the levels of hierarchy, i.e. time transitions must synchronise in all the levels of hierarchy of a hierarchical automaton.

- **State refinement operator** [15]. State refinement is a binary operator on process algebraic processes written as $p[q]$ where $p, q$ are arbitrary process terms. Informally, it means that $p$ is a state with the substructure $q$. In other words, a location of an automaton is allowed to contain another automaton
representing its substructure. Furthermore, it was also stated [15] that the above way of introducing hierarchy is compatible with the interleaving models of concurrency. Thus, the present work is motivated by the work carried out in [15], even though the basic entity in our formalism is an automaton rather than an action.

The semantics of the hCIF formalism is a hybrid transition system (HTS) [5] constructed using structural operational semantics (SOS) [12]. The goal of this paper is to show how the semantics of hierarchical automata can be defined using SOS rules, in a compositional way. We do so without introducing the complexity of state tree structures in the formalism, and in such a way that the semantics of other CIF operators remain unaffected. As an additional result, we define an algorithm for eliminating hierarchy, which enable us to reuse existing tools for implementing an hCIF simulator.

The remainder of this paper is organized as follows. First, the subset of CIF, which is extended with hierarchy, is presented in Section 2. Once the basic concepts are introduced, Section 3 introduces the syntactical extensions that are needed to add support for hierarchy in CIF, and we discuss the design-decisions that led to such extensions. The formal semantics of hCIF is presented in Section 4, where we illustrate how the concept of hierarchy can be defined in a compositional and recursive way. In Section 5, we give a procedure to eliminate the hierarchy from a hCIF model. Finally, in Section 6 we make some conclusive remarks and discuss future work.

2 Introduction to CIF

This section presents the syntax and semantics of a subset of CIF [3], because presenting the full syntax and semantics would distract too much from the message of the paper. Still, we have included all the aspects of the CIF that were involved in the design-decisions we made when defining hierarchy.

As was mentioned in the introduction, the CIF language is based on hybrid automata, which model a combination of computational and physical behavior of a system by mixing automata theory with the theory of algebraic differential equations. A key feature of the CIF is that it provides a structural operational semantics for such atomic (hybrid) automata, which makes the definition of more complex compositions of automata easier. The only compositions defined in this paper will be parallel composition and hierarchy, but in [3] many more are described.

Informally, a basic CIF automaton is shown in Figure 1 that models the dynamics of a thermostat in a room. This thermostat can be in one of two computational states. It is either off, or on. This is reflected by the two circles (called locations) labeled Off and On. Next to the label, the locations also contain equations (called time-can-progress predicates or tcp-predicates) that model the physical behavior of the system while it is in this computational state. In our example, the temperature $T$ will behave according to the differential equation $\dot{T} = -T + 15$ when the thermostat is off and according to $\dot{T} = -T + 25$ when the thermostat is on. The dotted version of variables, like $\dot{T}$ in the example, are used for modeling derivatives. The execution of a calculation generally results in a change of location, which is modeled by an arrow (called edge) from one location to another. Edges are labeled by actions (in our example switch-on and switch-off) that may be used to synchronize the behavior of automata in a composition (not shown here). Furthermore, they contain a predicate (called guard) that determines under which condition an action can be executed, and a predicate (called reset) that determines whether there is a change in any of the model variables. In case of the thermostat, the switch-on action can only be executed if the temperature $T$ is lower than 20, and the action results in a change of the variable $n$ to $n + 1$, which counts the number of times the thermostat is switched on. Notation $n^+$ is used to refer to the value of $n$ after the execution of the action. The guard $n \leq 1000$ disables the switch-off action, modeling that the thermostat breaks down
Hierarchical states in the Compositional Interchange Format

after a thousand switches and leaves the room hot. Every location contains an initialization predicate, which determines whether execution can start in that location. In the example, initially the thermostat is switched off, the temperature in the room is \( T = 25 \) and the counter is set to \( n = 0 \). This is modeled by an incoming edge in the Off location without any origin, labeled by the predicate \( T = 25 \land n = 0 \). The absence of incoming arrows on location Off denotes that the initialization predicate is false (which means that execution cannot start on that location).

Formally, the locations of an atomic CIF automata are taken from the set \( \mathcal{L} \). Actions belong to the set \( \mathcal{A} \). We distinguish the following types of variables: regular variables, denoted by the set \( \mathcal{V} \); the dotted versions of those variables, which belong to the set \( \dot{\mathcal{V}} \equiv \{ \dot{x} \mid x \in \mathcal{V} \} \); and step variables, which belong to the set \( \{ x^+ \mid x \in \mathcal{V} \cup \dot{\mathcal{V}} \} \). Furthermore, the variables can be classified according their continuous evolution (i.e. how their values change during time delays). In particular, we distinguish between discrete variables (\( n \) in the previous example), whose values remain constant during time delays, and the value of their dotted versions are always 0; and continuous variables (\( T \) in the previous example), whose values evolve as a continuous function of time during delays, and whose dotted versions represent their derivatives. Variables are constrained by differential algebraic equations and we implement them as predicates. The values of the variables belong to the set \( \Lambda \) that contains, among else, the sets \( \mathbb{B}, \mathbb{R}, \) and \( \mathbb{C} \). Guards, tcp and initialization predicates, and reset predicates are taken from the sets \( \mathcal{P}_g, \mathcal{P}_t, \) and \( \mathcal{P}_r \), respectively.

The exact syntax and semantics of predicates are left as a parameter of our theory, as we are not interested in the computational aspects of CIF in this paper. In the examples presented here, and in the tool implementations of CIF, \( \mathcal{P}_g \), \( \mathcal{P}_t \), and \( \mathcal{P}_r \) are terms of the language of predicate logic [13], where for \( \mathcal{P}_g \) and \( \mathcal{P}_t \) the variables are taken from the set \( \mathcal{V} \cup \dot{\mathcal{V}} \), and for \( \mathcal{P}_r \) the variables are taken from the set \( \mathcal{V} \cup \dot{\mathcal{V}} \cup \{ x^+ \mid x \in \mathcal{V} \cup \dot{\mathcal{V}} \} \).

Given these preliminaries, an atomic automaton can be defined as follows.

**Definition 2.1** (Atomic automaton). An atomic automaton is a tuple \((V, \text{init}, \text{tcp}, E)\) with a set of locations \( V \subseteq \mathcal{L} \); initial and time-can-progress predicates \( \text{init}, \text{tcp} : V \to \mathcal{P}_t \); and a set of edges \( E \subseteq V \times \mathcal{P}_g \times \mathcal{A} \times \mathcal{P}_r \times \mathcal{L} \).

We use symbol \( \mathcal{M} \) to refer to the set of all atomic automata. Atomic automata, as the one shown before, can be used to build more complex models by using the parallel composition operator. CIF includes more operators, but we do not discuss them in this paper. Throughout this work, we use the term composition to refer either to an atomic automata, or to a parallel composition of automata, which is denoted as \( p \parallel q \), for compositions \( p \) and \( q \), where the set \( S \subseteq \mathcal{A} \) is the set of actions that must be executed synchronously in both automata. The set \( \mathcal{E} \) contains all \( hCIF \subseteq \) compositions, and is formally defined in Section 3.

It is not possible to present the formal semantics of CIF in this paper due to the lack of space. However, if an automaton has no hierarchy (i.e. no location contains a composition), the rules presented here match those of CIF.

After introducing the base language, we are ready to show how it can be extended with hierarchy.

![Figure 1: A model of thermostat.](image-url)
3 Adding hierarchy to CIF

In this section, we show the syntactical extensions that are needed for adding hierarchy to CIF, explaining why every new element is required according to the design-decisions we have taken.

An automaton is said to be hierarchical if it contains a composition in at least one of its locations. Extending CIF with hierarchy is easy to achieve with the addition of a hierarchy function $h$ to the elements of an automaton, such that $h(v)$ returns the composition contained in $v$, which we refer to as substructure. We find this hierarchy function to be more suitable for the automata theoretic framework of CIF than the state refinement operator \[15\], which was conceived for process algebraic setting. This is because, unlike in process algebra, the development of the CIF is aimed at modeling convenience rather than at finding the smallest representation of a given construct.

As an example, suppose we want to extend the model of the thermostat presented in Figure 1, so that it is only switched off after a certain time has elapsed (which allows the room to be heated up). Having hierarchy, a refined model of the thermostat can be elaborated as in Figure 2. We define the hierarchy function $h$ such that $\text{Off} \notin \text{dom}(h)$; and $h(\text{On})$ returns the automaton shown at the bottom in Figure 2, which initially sets up a clock $c$ (continuous variable), and it waits until the timer expires $\Delta \leq c$ (where $0 < \Delta$) to generate the event done.

An important thing to note about the substructure in figure 2 is that one of the states contains an outgoing arrow that leads nowhere. The predicate on this arrow is called a termination predicate and it is taken, just like tcp and initialization predicates, from the set $P_t$.

Termination predicates were not part of the atomic automata of CIF before, even though they are common in the general theory of automata. Our reason for adding them is that we need a mechanism to decide when a substructure hands over the control of the execution flow to the superstate to which it belongs.

In most hierarchical formalisms, such as \[16\] and \[10\], the actions enabled in the super-automaton are executed regardless of the state of the substructure. The example in Figure 2, however, illustrates that a more general approach, in which the substructure has control over the superstructure, may be useful. The example actually uses the fact that the substructure has control over the superstructure to restrict the behavior of the thermostat in such a way that it is forced to stay in the On location for a certain time. The termination mechanism that we have chosen, originates from our desire to be able to (partially) translate sequential function charts \[2\] to CIF. In that formalism, termination is used as the standard mechanism...
to pass control from one chart to the next.

Another mechanism that we need when dealing with hierarchy, is a way to keep track of the so-called *active location* that an automaton is in. Admittedly, the active location of an automaton is a semantic concept, used to keep track of the current state when describing the dynamics of an automaton, and it does not belong in a section describing the syntax. However, in the structured operational semantics of the CIF, the states are formed by pairs of syntactic descriptions and valuations of variables (see section 4 for details), and our solution that deals with the semantic problem of keeping track of the active location, makes use of an auxiliary syntactic construct that one should not use while modeling, but that formally is part of the syntax of CIF.

In the non-hierarchical CIF, the active location of an automaton is fully determined by its initialization predicate. As a result, a state-change in the semantics is modeled by changing this initialization predicate. In the hierarchical CIF, however, changing the initialization predicates of a substructure means changing the hierarchy function $h$, and this causes so-called history retention; i.e. if a substructure is terminated and the automaton gets back to it later, the substructure will restart where it was ended previously, because $h$ is still in its altered form. From the example in Figure 2, one can see that this is not always desirable. History retention in that example, would make that the thermostat is only forced to linger in the *On* state the first time it is entered. In subsequent visits, the substructure would already be in its terminating state immediately.

The semantics we would like to give to hierarchical CIF, is that a substructure is restarted every time from one of its (original) initial states. Our solution for the history retention problem, is to introduce an auxiliary composition operator $p : \alpha$ which should be read as $\alpha$ is currently in the substructure $p$. Here, $p$ is an arbitrary composite automaton and $\alpha$ is an atomic hierarchical automaton. Next, the initialization predicate of $\alpha$ can be used to model the active location of the super-automaton $\alpha$, while the initialization predicate(s) of (the components) $p$ are used to model the active location of the active substructure. We found that the use of this auxiliary operator greatly simplifies the structured operational semantics of hierarchical CIF.

Now that we have informally introduced all the syntactic elements needed for extending CIF with hierarchy, we define the syntax of hierarchical automata as follows.

**Definition 3.1.** An atomic hierarchical hybrid automaton is a tuple $(V, \text{init}, \text{tcp}, E, \text{term}, h)$ where, $(V, \text{init}, \text{tcp}, E)$ is an atomic CIF automaton, $\text{term} : V \rightarrow \mathcal{P}$ is a function that associate to each location a predicate describing the conditions under which a location is final, and $h : V \rightarrow \mathcal{C}$ is a (partial) hierarchy function that maps each location with a composite automaton. The set of composite automata $\mathcal{C}$ in hCIF is recursively defined by the following grammar $\mathcal{C} ::= \alpha \mid \mathcal{C} \mid \mathcal{C} \parallel S \mathcal{C}$.

Henceforth, we use Greek letters $\alpha$ and $\alpha'$ to indicate an atomic hierarchical hybrid automaton and Roman letters $p, q, p'$, and $q'$ to indicate any composite automaton in $\mathcal{C}$.

## 4 Formal semantics of hCIF

In this section we illustrate how the concept of hierarchy can be defined in a compositional way, without introducing additional concepts of tree-structures present in the statechart formalism. First, the semantic framework is set up, and then the SOS rules are presented.

The semantics of hCIF compositions (and CIF) is given in terms of SOS rules, which induce hybrid transition systems (HTS) [5]. The states of the HTS are of the form $(p, \sigma)$, where $p \in \mathcal{C}$ is a composition and $\sigma : V \cup \hat{V} \rightarrow \Lambda$ is a function, called *valuation*, which assigns values to the variables. Valuations capture the phenomenon of discrete change in the values of variables caused by the execution of actions
in an automaton. We denote the set of all valuations as \( \Sigma \). There are three kind of transition in the HTS, namely, action transitions, environment transitions, and time transitions. We describe them in detail next.

Action transition are of the form \( \langle p, \sigma \rangle \xrightarrow{a} \langle p', \sigma' \rangle \), and they model the execution of action \( a \) by process \( p \) in an initial valuation \( \sigma \), which changes process \( p \) into \( p' \) and results in a valuation \( \sigma' \).

Environment transitions are of the form \( \langle p, \sigma \rangle \xrightarrow{b} \langle p', \sigma' \rangle \), and in the full CIF language, they are used to model which possible behavior of the environment is consistent with that of the composition \( p \), but cannot be executed by the component itself. In the restricted language hCIF, the function of the environment transitions is to indicate that a composition \( p \) can initialize to become a composition \( p' \) in which an active location is fixed for each (active) substructure. Furthermore, the boolean \( b \) indicates whether the initialized substructure can terminate, and thus give back the control over actions to their environment.

Time transitions are of the form \( \langle p, \sigma \rangle \xrightarrow{\rho, \theta, \omega} \langle p', \sigma' \rangle \), and they model the passage of time in composition \( p \), in an initial valuation \( \sigma \), which results in a composition \( p' \) and valuation \( \sigma' \). The relation between \( p \) and \( p' \) is the same as for environment transitions. Function \( \rho : \mathbb{T} \to \Sigma \) is called variable trajectory, and it models the evolution of variables during the time delay. For each time point \( s \in \text{dom}(\rho) \), and for each variable \( x \in \mathcal{V} \cup \mathcal{V}' \), the function application \( \rho(s)(x) \) yields the value of variable \( x \) at time \( s \). Function \( \theta : \mathbb{T} \to 2^\mathcal{V} \) is called guard trajectory, and it models the evolution of enabled actions during time delays. For each time point \( s \in \text{dom}(\theta) \), the function application \( \theta(s) \) yields the set of enabled actions of composition \( p \) at time \( s \). Lastly, function \( \omega \) is called termination trajectory, and it models the evolution of termination during time delays: for each time point \( s \in \text{dom}(\omega) \), composition \( p' \) is terminating at time \( s \) if and only if \( \omega(s) \). For all time transition \( \text{dom}(\rho) = [0, t] \), for some time point \( t \in \mathbb{T} \), and \( \text{dom}(\rho) = \text{dom}(\theta) = \text{dom}(\omega) \).

Guard trajectories were shown to be necessary for the definition of urgency and variable abstraction in CIF and other hybrid formalisms \[4, 18\]. Termination trajectories allow us to keep track of the possibility of termination over time, and they are essential for constructing the guard trajectories in the rules. Even though these concepts are not necessary for giving semantics to hCIF, they allow us to solve the problem of supporting urgency and variable abstraction in a hierarchical setting, and thus our approach can be extended to the whole CIF without modifying the rules.

Even though predicates are abstract entities, we assume that there is a satisfaction relation \( \sigma \models e \) is defined, which expresses that predicate \( e \in \mathcal{P}_t \cup \mathcal{P}_p \cup \mathcal{P}_r \) is satisfied (i.e. it is true) in valuation \( \sigma \). For predicate logic, this relation can be defined in a standard way (see \[13\] for example). For a valuation \( \sigma \), we define \( \sigma^+ \triangleq \{(v^+, c) | (v, c) \in \sigma\} \).

Definition \[4.1\] formalizes the hybrid transition system induced by the SOS rules presented in the next sections.

**Definition 4.1.** A hybrid transition system (HTS) is a six-tuple of the form \( (Q, \mathcal{A}, \rightarrow, \longrightarrow, \cdots) \) where, \( Q \triangleq \mathcal{V} \times \Sigma, \rightarrow \subseteq Q \times \mathcal{A} \times Q, \longrightarrow \subseteq Q \times (\mathbb{T} \to \Sigma) \times (\mathbb{T} \to 2^\mathcal{V}) \times (\mathbb{T} \to \mathbb{B}) \times Q \), and \( \cdots \subseteq Q \times \mathbb{B} \times Q \).

### 4.1 Hierarchical hybrid automaton

In this section, we give semantics to hierarchical hybrid automata. We use notation \( \alpha \) to refer to an atomic automaton \( (V, \text{init}, \text{tcp}, E, \text{term}, h) \), and \( \alpha[v] \) to refer to the automaton \( (V, \text{id}_v, \text{tcp}, E, \text{term}, h) \), where \( \text{id}_v(w) \triangleq v = w \).

In absence of hierarchy, an atomic automaton \( \alpha \) can perform an action in a location \( v \) and initial valuation \( \sigma \) if there is an edge \((v, g, a, r, v')\) such that the following conditions hold:
1. Location $v$ is active ($\sigma \models \text{init}(v)$).

2. Guard $g$ holds ($\sigma \models g$).

3. It is possible to find a new valuation $\sigma'$ such that the reset predicate is satisfied in valuation $\sigma \cup \sigma'^+ \quad (\sigma \cup \sigma'^+ \models r)$. We do not write $\sigma' \models r$ since, in general, $r$ refers to the next values of variables, which are contained in $\sigma'^+$.

The above conditions are summarized in the term $\sigma, \sigma' \models_a (v, g, a, r, v')$, which is syntactically equivalent to condition $(v, g, a, r, v') \in E \land \sigma \models \text{init}(v) \land \sigma \models g \land \sigma'^+ \cup \sigma \models r$.

Rule 1 describes what is semantically involved with the addition of hierarchy. Firstly, it is necessary to check that the substructure of the initial location, if any, is terminating (condition $\langle h(v), \sigma \rangle \models \text{true} \land \sigma \models \text{init}(v) \land \sigma \models g \land \sigma'^+ \cup \sigma \models r$). Note that the choice of selecting active locations of substructure $h(v')$ is made upon entering location $v'$. Example 1 illustrates this phenomenon, which we call eager choice.

$$
\begin{align*}
\sigma, \sigma' & \models_a (v, g, a, r, v'),
\left(\langle h(v), \sigma \rangle \models \text{true} \land \sigma \models \text{init}(v) \land \sigma \models g \land \sigma'^+ \cup \sigma \models r\right)

\sigma', v' & \in \text{dom}(h),
\left(\langle h(v'), \sigma' \rangle \models h \land \sigma' \models \text{init}(v') \land \sigma' \models g \land \sigma'^+ \cup \sigma' \models r\right)

\langle \alpha, \sigma \rangle & \xrightarrow{a} \langle q : \alpha[v'], \sigma' \rangle
\end{align*}
$$

**Example 1.** Consider the composite automaton shown in Figure 3a in which the tcp function gives the predicate true for all the vertices of the automaton. The idea behind eager choice, is that after the execution of the action $a$, an initial state of the substructure is picked immediately. This can only result in the left state of the substructure to be picked, because of the value of $x$ is set to 1 during the execution of $a$. Hence, the action $c$ in the substructure will never be executed. The resulting transition system generated by the SOS is shown in Figure 3b, where the states are depicted as circles and their components are not shown. End of Example.

![Figure 3: (a) Automaton with two possible initial states. (b) Resulting HTS.](image)

Rule 1 requires as a condition that there is an active substructure in the target location $v' \in \text{dom}(h)$. If this is not the case then no active substructure is prefixed to $\alpha[v]$, as expressed by Rule 2.

$$
\begin{align*}
\sigma, \sigma' & \models_a (v, g, a, r, v'), \quad v' \notin \text{dom}(h),
\left(\langle h(v), \sigma \rangle \models \text{true} \land \sigma \models \text{init}(v) \land \sigma \models g \land \sigma'^+ \cup \sigma \models r\right)

\langle \alpha, \sigma \rangle & \xrightarrow{a} \langle q : \alpha[v'], \sigma' \rangle
\end{align*}
$$
In a hierarchical setting, actions in an automaton can be generated by the substructure of an active location. Rule 3 formalizes this. Note that in the conclusion, $p : \alpha[v]$ reflects the fact that an initial location is chosen in a hierarchical structure if the substructure performs an action.

$$
\sigma \models \text{init}(v), v \in \text{dom}(h), \langle h(v), \sigma \rangle \xrightarrow{a} \langle p, \sigma' \rangle
$$

In CIF, a time delay is possible in an active location $v$ if there exists a trajectory $\rho$ such that the tcp predicate is satisfied in $[0, r)$. Henceforth, we will use $t, \rho \models_{a} \langle \text{init}(v), \text{tcp}(v) \rangle$ as an abbreviation of the predicate $\rho(0) \models \text{init}(v) \land \text{dom}(\rho) = [0, t) \land 0 \leq t \land \forall s \in [0, t), [\rho(s) \models \text{tcp}(v)]$.

For time delays, in hCIF the substructure must perform a time transition with the same trajectory, and we consider conjunction of all the tcp predicates of all the active locations of an automaton. In this way time passes in an automaton, and also in all of its contained substructures. This is, an automaton and its substructure synchronize in the time delays. In the complete extension of CIF with hierarchy, a similar approach is taken for invariants. Rule 4 models this, where $\forall x \in [0,t]. \omega(x) = \omega_{0}(x) \land \rho(x) \models \text{term}(v)$, and $\forall x \in [0,t]. \theta(x) = \theta_{0}(x) \cup \{ a \mid (v, g, a, r, v') \in E \land \rho(s) \models g \land \omega_{0}(s) \}$. The guard trajectory $\theta$ as well as the termination trajectory $\omega$ are constructed by using the corresponding trajectories generated by the time transition in the substructure.

We found that this approach is the simple and intuitive: substructures are part of the whole structure, and it is strange from the modeling point of view to have a system in which time can “freeze” for certain parts of the system. Furthermore, since there are several active locations in different levels of hierarchy, it is not clear which one to choose to perform the time delays. The example of the thermostat with hierarchy, depicted in Figure 2, shows the convenience of this decision, since we want the clock to advance while the room heats up.

$$
t, \rho \models_{a} \langle \text{init}(v), \text{tcp}(v) \rangle, \langle h(v), \rho(0) \rangle \xrightarrow{\rho, \theta_{0}, \omega_{0}} \langle p, \rho(r) \rangle
$$

The following example illustrates the need for having termination trajectories to capture properly the set of enabled actions during time delays.

**Example 2.** Consider the automaton shown in Figure 4a and assume $1 < k_0 < k_1$. Then the set of enabled actions at the active location will depend on the function $e^x$ and this set is illustrated in Figure 4b. In other words, if $0 \leq x < \ln(k_0)$ then the set of enabled actions is $\{ a \}$. If $\ln(k_0) \leq x < \ln(k_1)$ then the set of enabled actions is $\{ a, b \}$. And if $x \geq \ln(k_1)$ then the set of enabled actions is $\{ a \}$. **End of Example.**
Rule 5 deals with the case when an initial location \( v \) does not contain a substructure, where \( \text{dom}(\omega) = \text{dom}(\rho) \) and \( \forall s \in [0,t]. \omega(s) = \rho(s) \models \text{term}(v) \), and \( \forall s \in [0,t]. \theta(s) = \{ a \mid (v,g,a,r,v') \in E \land \rho(s) \models g \} \).

\[
\begin{align*}
\frac{t, \rho \models \alpha \text{(init}(v) , \text{tcp}(v)) , v \not\in \text{dom}(h)}{\langle \alpha, \rho(0) \rangle \xrightarrow{\rho, \theta, \omega} \langle \alpha[v], \rho(t) \rangle}
\end{align*}
\] (5)

In CIF, if an automaton performs an environment transition then an unique active location is chosen. When hierarchy is incorporated, the substructure is initialized as well. This is expressed by Rule 6. The initialized composition \( p \) becomes the active substructure of \( \alpha[v] \), and the automaton is terminating if the location and the active substructure are. Rule 7 deals with the case where there is no substructure.

\[
\begin{align*}
\frac{\sigma \models \text{init}(v), \langle h(v), \sigma \rangle}{\langle \alpha, \sigma \rangle \xrightarrow{\sigma \models \text{term}(v) \land b} \langle p, \sigma' \rangle} \quad \frac{\sigma \models \text{init}(v), v \not\in \text{dom}(h)}{\langle \alpha, \sigma \rangle \xrightarrow{\sigma \models \text{term}(v)} \langle \alpha[v], \sigma' \rangle}
\end{align*}
\] (6) (7)

### 4.2 Automaton postfix operator

We now define the SOS rules for the automaton postfix operator, which helps in defining the overall behavior of a hierarchical automaton.

Intuitively, the composition \( p : \alpha \) means that composition \( p \) is the active substructure of some initial location \( v \in V \) in the automaton \( \alpha \). Note, that whenever the automaton postfix is only used as an auxiliary operator, this initial location will always be uniquely specified.

If we now look at the structure of a state \( (p, \sigma) \) in the hybrid transition system, it becomes clear how the postfix operator helps us to mimick the state-tree structures used in the semantics of statecharts [8]. Figure 5 shows that a composition \( p \) in essence is a tree, where the postfix operator represents the edges of the tree and the parallel compositions represents the branching. The root of this tree is the active location of the automaton we described, while the leaves are the active substructures where the control over actions currently lies. Indeed, an informal comparison of our semantics to that of statecharts suggest that the AND-superstates of statecharts are represented as (asynchronous) parallel compositions \( ||0 \), while OR-superstates are represented by having multiple locations for which the initialization predicate holds.

\[
\begin{align*}
\text{Figure 5: Relation between automaton postfix operator and state tree structures}
\end{align*}
\]

The semantics of \( p : \alpha \) is reminiscent of the sequential composition of untimed process algebra, i.e. the composite automaton \( p \) in \( p : \alpha \) will perform action transitions until it is terminating, after that the automaton \( \alpha \) can perform its action transitions. The difference between the sequential composition
operator and the automaton postfix operator is due to the difference in the passage of time caused by these operators. In the automaton postfix operator, the passage of time is synchronized between the first and second component, whereas in sequential composition the passage of time is not synchronized. There the second component waits for termination of the first, in a similar way as with action transitions.

All the rules presented here are similar to the those presented in the previous section. The difference lies in that in the rules of this section function $h$ is not considered, since there is an active substructure $p$ in the target state of every transitions appearing in the conclusions. Rule $8$ models the action transition when the substructure is terminating.

$$\sigma, \sigma' = (v, g, a, r, v'), \langle p, \sigma \rangle \longrightarrow \langle p', \sigma \rangle \longrightarrow \langle q, \sigma' \rangle \quad (8)$$

Rule $9$ also models the action transition generated from a postfix operator when the substructure is terminating and the target location $v'$ does not contain a substructure. Rule $10$ models the action transition which is a result of the execution of the substructure.

$$\sigma, \sigma' = (v, g, a, r, v'), \langle p, \sigma \rangle \longrightarrow \langle p', \sigma \rangle, v' \notin \text{dom}(h) \quad (9) \quad \langle p, \sigma \rangle \xrightarrow{\alpha} \langle q, \sigma' \rangle \quad (10)$$

Finally, Rule $11$ models the passage of time in an automaton postfix such that the timed transitions are synchronized in every level of hierarchy $p : \alpha$, where, $\text{dom}(\omega) = \text{dom}(\rho)$, $\text{dom}(\theta) = \text{dom}(\rho)$ and $\forall s \in [0, t], \omega(s) = \omega_0(s) \land \rho(s) = \text{term}(v)$, and $\forall s \in [0, t], \theta(s) = \theta_0(s) \cup \{a \mid (v, g, a, r, v) \in E \land \rho(s) \models g \land \omega_0(s)\}$.

$$t, \rho \models (\text{init}(v), \text{tcp}(v)), \langle p, \rho(0) \rangle \xrightarrow{\rho, \theta_0, \omega_0} \langle p', \rho(t) \rangle \quad (11)$$

Rule $12$ models the execution of environment transition in an automaton postfix.

$$\langle p, \sigma \rangle \xrightarrow{b} \langle p', \sigma' \rangle \quad (12)$$

$$\langle p : \alpha, \sigma \rangle \xrightarrow{\text{term}(v) \land b} \langle p' : \alpha[v], \sigma' \rangle$$

### 4.3 Parallel composition operator

The parallel composition operator allows synchronisation of equally labeled action transitions between any two components that are specified by the synchronisation set $S \subseteq \mathcal{A}$. Rule $13$ models this fact.

$$\langle p, \sigma \rangle \xrightarrow{a} \langle p', \sigma' \rangle, \langle q, \sigma \rangle \xrightarrow{a} \langle q', \sigma' \rangle, a \in S \quad (13)$$

Rule $14$ models the interleaving of action transitions that do not belong to the synchronisation set $S$. Note the presence of environment transition in the premise of the following rule. This allows the other component (which does not perform an action transition, in the following case $q$) to get initialised.

$$\langle p \parallel s q, \sigma \rangle \xrightarrow{a} \langle p' \parallel s q', \sigma' \rangle$$

$$\langle q \parallel s p, \sigma \rangle \xrightarrow{a} \langle q' \parallel s p', \sigma' \rangle$$
Hierarchical states in the Compositional Interchange Format

\[ \langle p, \sigma \rangle \overset{a}{\rightarrow} \langle p', \sigma' \rangle, \langle q, \sigma \rangle \overset{b}{\rightarrow} \langle q', \sigma' \rangle, a \not\in S \]

(14)

\[ \langle p \parallel s q, \sigma \rangle \overset{a}{\rightarrow} \langle p' \parallel s q', \sigma' \rangle, \langle q \parallel s p, \sigma \rangle \overset{a}{\rightarrow} \langle q' \parallel s p', \sigma' \rangle \]

In a parallel composition, time can pass independently in each component, as it can be seen in Rule [15] where \( \forall s \in [0, t], [\theta_0(s) = (\theta_0(s) \cap \theta_1(s)) \cup (\theta_0(s) \setminus S) \cup (\theta_1(s) \setminus S)] \) and \( \forall \sigma \in [0, t], [\omega_0(s) = \omega_0(s) \land \omega_1(s)] \). The guard trajectory is constructed (same as in CIF) from the guard trajectories of the composite automata interleaving in the parallel composition: a given time point \( s \) is an event of \( p \) if it is enabled in \( p \) and \( q \) (regardless of whether the action is in \( S \)), or if it is enabled in \( p \) or \( q \) and is not synchronizing in the other component.

The termination trajectory in a parallel composition at a given time point \( s \) is the conjunction of the termination trajectories of the respective components at the same time instant \( s \).

\[ \langle p, \rho(0) \rangle \overset{p, \theta_0, \omega_0}{\rightarrow} \langle p', \rho(t) \rangle, \langle q, \rho(0) \rangle \overset{p, \theta_1, \omega_0}{\rightarrow} \langle q', \rho(t) \rangle \]

(15)

\[ \langle p \parallel s q, \rho(0) \rangle \overset{p, \theta_0, \omega_0}{\rightarrow} \langle p' \parallel s q', \rho(t) \rangle, \langle q \parallel s p, \rho(0) \rangle \overset{p, \theta_1, \omega_0}{\rightarrow} \langle q' \parallel s p', \rho(t) \rangle \]

The initialization of a parallel composition (Rule [16]) is the initialization of its components. The termination predicate in the parallel composition is the conjunction of the termination predicates of the respective components.

\[ \langle p, \sigma \rangle \overset{b_0}{\rightarrow} \langle p', \sigma' \rangle, \langle q, \sigma \rangle \overset{b_1}{\rightarrow} \langle q', \sigma' \rangle \]

(16)

\[ \langle p \parallel s q, \sigma \rangle \overset{b_0}{\rightarrow} \langle p' \parallel s q', \sigma' \rangle, \langle q \parallel s p, \sigma \rangle \overset{b_1}{\rightarrow} \langle q' \parallel s p', \sigma' \rangle \]

4.4 Stateless bisimulation

It is clear from the definition of a HTS (Definition [14]) that a state in a transition system consists of a transition part (a behavioural part) and a data part (valuation). Furthermore, we know that the stateless bisimulation is the most robust equivalence for the transition systems whose states contains data [11]. This subsection shows that the semantics of hCIF\(^C\) is compositional with respect to stateless bisimulation [11], i.e. stateless bisimulation is a congruence for all operators of hCIF\(^C\).

Definition 4.2. A symmetric relation \( R \subseteq \mathcal{E} \times \mathcal{E} \) is called a stateless bisimulation [11] relation iff the following transfer conditions hold.

\[ \forall p, p', \sigma, \sigma', a, q. \left[ \langle p, \sigma \rangle \overset{a}{\rightarrow} \langle p', \sigma' \rangle \land \langle p, q \rangle \in R \Rightarrow \exists q'. \left[ \langle q, \sigma \rangle \overset{a}{\rightarrow} \langle q', \sigma' \rangle \land \langle p', q' \rangle \in R \right] \right] . \]

\[ \forall p, p', \sigma, \sigma', \rho, \theta, \omega, q. \left[ \langle p, \sigma \rangle \overset{\rho, \theta, \omega}{\rightarrow} \langle p', \sigma' \rangle \land \langle p, q \rangle \in R \Rightarrow \exists q'. \left[ \langle q, \sigma \rangle \overset{\rho, \theta, \omega}{\rightarrow} \langle q', \sigma' \rangle \land \langle p', q' \rangle \in R \right] \right] . \]

\[ \forall p, p', \sigma, \sigma', b, q. \left[ \langle p, \sigma \rangle \overset{b}{\rightarrow} \langle p', \sigma' \rangle \land \langle p, q \rangle \in R \Rightarrow \exists q'. \left[ \langle q, \sigma \rangle \overset{b}{\rightarrow} \langle q', \sigma' \rangle \land \langle p', q' \rangle \in R \right] \right] . \]

Two composite automata \( p, q \) are said to be stateless bisimilar (denoted \( p \leftrightarrow s q \)) iff there exists a stateless bisimulation relation \( R \) such that \( \langle p, q \rangle \in R \).
Theorem 4.3. Stateless bisimulation is a congruence for all the constructs of hCIF$^\subset$.

Proof. The SOS rules of hCIF$^\subset$ are in the process-tyft format, which guarantees the congruence for stateless bisimilarity [11].

5 Elimination of hierarchy

In this section, we present a technique that converts a hierarchical automaton into an automaton in which the hierarchical function $h$ is empty, such that they are stateless bisimilar. Such techniques in general, are known as linearization or elimination of operators [17]. The advantage of flattening a hierarchical automaton is that it allows the reuse of existing tools like simulators of the CIF language, which are only developed for flat automata.

Definition 5.1. The depth $D(p)$ of an composite automaton $p \in \mathcal{C}$ is recursively defined to be the $1 + \max_{v \in \text{dom}(h)} D(h(v))$ when $p$ is a hierarchical automaton of the form $(V, \text{init}, tcp, E, term, h)$, and is defined as $\max(D(q), D(r))$ whenever $p = q \parallel s r$. An automaton is called well founded whenever its depth is defined. An automaton of depth 1 (i.e. with $\text{dom}(h) = \emptyset$) is also called flat automaton.

Suppose that we have a procedure $\mathcal{E}$ that turns any composition of flat automata into a stateless bisimilar flat automaton. In particular, suppose that $\mathcal{E}(\alpha \parallel s \alpha') \leftrightarrow_s \alpha \parallel s \alpha'$ whenever $\alpha$ and $\alpha'$ are flat automata. Then, we can lift this procedure to any well-founded composite automaton $p \in \mathcal{C}$, by first applying it to all components of $p$ before applying it to $p$ itself. We define $\mathcal{E}(p \parallel s q) = \mathcal{E}(\mathcal{E}(p) \parallel s \mathcal{E}(q))$ and $\mathcal{E}((V, \text{init}, tcp, E, \text{term}, h)) = \mathcal{E}((V, \text{init}, tcp, E, \text{term}, h))$ with $h(v) = h(v)$, for any $p, q$ and $(V, \text{init}, tcp, E, \text{term}, h)$ of depth greater than 2. Structural induction on the depth of the composite automaton, combined with the congruence obtained in theorem 4.3, then gives us $\mathcal{E}(p) \leftrightarrow_s p$ for all well-founded composite automata $p$.

Such a procedure is already known for all the usual operations of the CIF [17], and next, we will give it for hierarchical automata.

Definition 5.2. Let $\alpha \in \mathcal{C}$ be an automaton of depth 2 of the form $(V, \text{init}, tcp, E, \text{term}, h)$, such that $h(v)$ is a flat automaton for all $v \in \text{dom}(h)$. We define $\mathcal{E}(\alpha) = (\hat{V}, \text{init}, \text{tcp}, \hat{E}, \text{term}, \emptyset)$ where,

- The set $\hat{V}$ of locations of the flat automaton $\mathcal{E}(\alpha)$ is defined by:
  
  $$\hat{V} \triangleq \bigcup_{v \in V} \left\{ (v, w) \mid \begin{array}{ll} v \notin \text{dom}(h) \land w = \perp & \lor \\
  h(v) = (V', \text{init}', \text{tcp}', E', \text{term}', \emptyset) \land w \in V' \end{array} \right\}$$

- The predicate-functions $\hat{\Box}$, with $\Box \in \{\text{init}, \text{tcp}, \text{term}\}$, are defined for each $\hat{v} \in \hat{V}$ by:
  
  $$\hat{\Box}(\hat{v}) \triangleq \begin{cases} \Box(v), & \text{if } \hat{v} = (v, \perp) \\
  \Box(v) \land \Box'(w), & \text{if } (\hat{v} = (v, w) \land h(v) = (V', \text{init}', \text{tcp}', E', \text{term}', \emptyset) \land w \in V') \end{cases}$$

- The edges $(\hat{v}_0, g', a, r', \hat{v}_1)$ of the flat automaton $\mathcal{E}(\alpha)$ are present (i.e. $(\hat{v}_0, g', a, r', \hat{v}_1) \in \hat{E}$) iff one of the following conditions hold:
  
  1. $(v_0, w_0) \land \hat{v}_1 = (v_1, w_1)$ for some $v_0, v_1 \in V$ such that $h(v_i) = (V'_i, \text{init}'_i, \text{tcp}'_i, E'_i, \text{term}'_i, \emptyset)$, $w_i \in V'_i$, for $i \in \{0, 1\}$ and

  $$(v_0, g, a, r, v_1) \in E \land g' = ((\text{term}'_0(w_0) \land g) \land r' = (r \land \text{init}'_1(w_1)^+) \land \text{true})$$

  Note that if $w_0 = \perp (w_1 = \perp)$ then by defining $\text{term}'_0(w_0) = \text{true} (\text{init}'_1(w_1) = \text{true})$ we get definitions for the simpler cases derived from the above one.
Hierarchical states in the Compositional Interchange Format

2. \( \dot{v}_0 = (v, w_0) \land \dot{v}_1 = (v, w_1) \) for some \( v_0 \in V \) such that

\[
h(v) = (V', \text{init}', \text{tcp}', E', \text{term}', \emptyset) \land w_0, w_1 \in V' \land (w_0, g', a, r', w_1) \in E'.
\]

Figure 5 shows the resulting automaton after applying the linearization procedure to the thermostat model extended with a clock (Figure 2).

Next we prove the correctness of our linearization procedure as defined in Definition 5.2.

**Theorem 5.3.** Let \( \alpha \in C \) be an automaton of depth 2 of the form \((V, \text{init}, \text{tcp}, E, \text{term}, h)\), such that \( h(v) \) is a flat automaton for all \( v \in \text{dom}(h) \), then \( \alpha \leftrightarrow s \mathcal{E}(\alpha) \).

**Proof.** Fix \( \hat{\alpha} = \mathcal{E}(\alpha) = (\hat{V}, \text{init}, \text{tcp}, \hat{E}, \text{term}, \emptyset) \). It is rather straightforward but tedious to verify that the relation \( R = \{(\alpha, \hat{\alpha}), (\alpha[v], \hat{\alpha}[(v, \perp)]), (h(v)[w] : \alpha[v], \hat{\alpha}[(v, w)]) \mid \forall \alpha \in V \land (v, w) \in \hat{V} \} \), is a witnessing stateless bisimulation. \( \square \)

## 6 Conclusions

In this paper we illustrated how to add hierarchy to a subset of the CIF (called hCIF \( ^{\subset} \)) in a compositional way, and we showed that the SOS rules of atomic entities can be modified without altering the rules of the CIF operators. Moreover, the rules are formulated in such a way that the addition of concepts such as urgency and invariants can be incorporated easily without altering the rules presented here. However, the usability of hCIF \( ^{\subset} \) is not yet investigated and the plan is to evaluate it by performing industrial case-studies within the context of the MULTIFORM project [11] after extending it with the remaining operators of the CIF. A procedure to eliminate hierarchy was given in order to be able to use the existing tools associated with CIF. Note that Definition 5.2 presented here is a relatively inefficient one to implement. It can be further optimized, for example, by disallowing the edges in the set \( \hat{E} \) that are never executed for any valuation.

As ongoing work, we are researching a branching version of stateless bisimulation for hybrid transition systems in order to handle \( \tau \) action as ‘invisible’. Using such a notion, it becomes possible to formalize when a stepwise refinement is a correct refinement of an abstract model. Thus we can attack the second aspect of stepwise refinement mentioned in the introduction.

**Acknowledgements:** The authors would like to thank Jos Baeten, Koos Roorda and Mohammad Mousavi, and the reviewers of SOS, for their constructive comments that very much helped to improve this paper.
This work has been performed as part of the Integrated Multi-formalism Tool Support for the Design of Networked Embedded Control Systems (MULTIFORM) project, supported by the Seventh Research Framework Programme of the European Commission. Grant agreement number: INFSO-ICT-224249.

References