INFORMATION AND COMPUTATION 79, 210–256 (1988)

Compositional Semantics for Real-Time Distributed Computing*

R. KOYMANS[†,‡,§]

Department of Mathematics and Computing Science, Eindhoven University of Technology, Den Dolech 2, P.O. Box 513, 5600 MB Eindhoven, The Netherlands

R. K. SHYAMASUNDAR[//]

NCSDCT, Tata Institute for Fundamental Research, Homi Bhaba Road, Bombay—400 005, India

W. P. DE ROEVER[‡] AND R. GERTH[‡]

Department of Mathematics and Computing Science, Eindhoven University of Technology, Den Dolech 2, P.O. Box 513, 5600 MB Eindhoven, The Netherlands

AND

S. ARUN-KUMAR

NCSDCT, Tata Institute for Fundamental Research, Homi Bhaba Road, Bombay-400 005, India

We give a compositional denotational semantics for a real-time distributed language, based on the linear history semantics for CSP of Francez *et al.* Concurrent execution is *not* modelled by interleaving but by an extension of the maximal parallelism model of Salwicki and Müldner, that allows for the modelling of transmission time for communications. The importance of constructing a semantics (and, in general, a proof theory) for real-time is stressed by such different sources as the problem of formalizing the real-time aspects of Ada and the elimination of errors in the real-time flight control software of the NASA space shuttle (*Comm. ACM* **27** (1984)).

* This paper is an extension of a preliminary version presented at the 1985 Logics of Programs Conference, Brooklyn, June 17–19, 1985. This research was done as an activity in the Dutch National Project Concurrency (Dutch acronym LPC).

[†] Supported by the Foundation for Computer Science Research in the Netherlands (SION) with financial aid from the Netherlands Organization for Scientific Research (NWO).

[‡] The author is currently working in and partially supported by ESPRIT project 937: Debugging and Specification of Ada Real-Time Embedded Systems (DESCARTES).

[§] Electronic mail address: mcvax!eutrc3!wsinronk.UUCP or WSDCRONK@HEITUE5.BITNET.

[//] Supported by a visitors grant from the Netherlands Organization for Scientific Research (NWO).

0890-5401/88 $3.00

1. INTRODUCTION

Although concurrency in programming has been seriously investigated for more than 25 years (Dijkstra, 1959), the specific problems of real-time have been the object of little theoretical reflection. Currently used real-time languages represent almost no evolution with respect to assembly language (Camerini, 1982). Consequently, no serious analysis of complexity, no design methodology, no standard for implementation, and no concept of portability exist for real-time languages.

The response to this has been the development of new real-time languages such as (1) Ada—developed for the military; (2) CHILL—within the context of telecommunication industries; and (3) Occam—which is even chip-implemented—for those interested in experimenting with structure. All of these are claimed to have been rigorously defined (Ada, 1983; Bjørner and Oest, 1980; Branquart, Louis, and Wodon, 1982; Occam, 1984). Yet their official standards lack any acceptable characterization of concurrency (with the exception of Occam), let alone of real-time (which is also lacking for Occam).

All these arguments emphasize the need to develop formal models for real-time concurrency, and, what is more important, to discover structuring methods which lead to hierarchical and modular development of real-time concurrent systems. Obviously, models based on interleaving, such as (Bernstein and Harter, 1981), can be immediately discarded as being unrealistic, since such models allow unbounded delay to be incurred between any two actions in a concurrent component.

A model such as SCCS (Milner, 1983), although an improvement by allowing truly concurrent activity, remains unsatisfactory because it either enforces complete synchronicity in executions (so that any communication must be performed immediately to circumvent deadlock) or it does not exclude interleaving (by using delay-operations). Petri-net theory remains a viable direction for discovering structuring methods, yet is still unsatisfactory because it does not incorporate (1) satisfactory verification methods for liveness properties, such as temporal logic has, or (2) (machine checkable) formalisms for representing (concurrently implemented) data structures. And certainly none of these models apply to real-time features of realistic programming languages such as Ada.

The present paper aims at providing a model of real-time concurrency which

— is *realistic* in the sense that concurrent actions can and will overlap in time unless prohibited by synchronization constraints, no unrealistic waiting of processors is modelled, and yet the many parameters involved in real-time behaviour are reflected by a corresponding

parametrization of our models (see Sections 9 and 10); it is based on Salwicki's notion of maximal concurrency (Salwicki and Müldner, 1981), discussed in Section 3;

— applies to programming languages for distributed computing such as Ada and Occam which are based on *synchronized communication* (for asynchronized communication as in CHILL, see Koymans, Vytopil, and de Roever (1983));

— implies a sound and relatively complete method for verification since it is *compositional*; we base ourselves in this respect on the method developed by Misra and Chandy (1981) and Zwiers (1985, 1988), and joint research together with Pnueli leading to the incorporation of maximal parallelism within the temporal framework of Barringer, Kuiper, and Pnueli (1984);

— meets the standard of rigour as provided by denotational semantics.

Some of these aspects are also covered by work of Zijlstra (1984) and G. Jones (1982).

We have developed a real-time variant of CSP (Hoare, 1978), called CSP-R, which allows the modelling of the essential Ada (1983) real-time features (see Appendix A). Our study of real-time distributed computing is carried by a subset of this language, Mini CSP-R (see Section 2). Extending our techniques to CSP-R introduces some notational complications, but is straightforward and is briefly discussed in Appendix A. In this paper we develop a denotational semantics for Mini CSP-R (in Section 7), stressing compositionality, based on the linear history semantics for CSP of Francez, Lehmann, and Pnueli (1984):

— the basic domain consists of non-empty *prefix-closed sets* of pairs of states and (finite) histories of communication assumptions leading to that state;

— the ordering on this domain is simply set-inclusion;

— the denotation for the parallel execution of two processes yields a denotation *in the same domain* for a new combined process replacing the original two (this makes the approach applicable to nested parallelism);

— the histories contain enough information to detect deadlock, eliminating the expectation states of Francez, Lehmann, and Pnueli (1984).

The basic domain and its interpretation is given in Section 6.

Histories are modelled as sequences of *bags* of communication assumption records as we allow truly concurrent actions: There is a clear operational difference between one process offering a particular communication capability and two (or more) processes, executing in parallel,

each offering the same capability. It is to model this distinction that we have to use bags instead of sets (see also Example 3 in Section 8).

The general notations and technical preliminaries for these concepts are defined in Section 5 which serves as a general reference point.

Real-time is modelled in the histories by relating the ith element of a history with the ith tick of a **conceptual** *global clock* (see Section 4).

There are two kinds of records for expressing communication assumptions in the histories:

— communication claims $\langle i, j, v \rangle$, modelling the execution of an I/O command: $\langle i, j, v \rangle$ claims that the value v is passed from process i (the sender) to process j (the receiver).

— no-match claims $\langle i, j \rangle$, modelling the absence of a possibility for the execution of an I/O command α (this means that there is no matching I/O command $\bar{\alpha}$ such that α and $\bar{\alpha}$ can be executed simultaneously): $\langle i, j \rangle$ claims that no value could be passed from process i (the sender) to process j (the receiver).

The combination of the communication assumption records $\langle i, j, v \rangle$ and $\langle i, j \rangle$ can be used to describe all possibe behaviours when executing an I/O command concerning communication from i to j: $\langle i, j, v \rangle$ claims that communication from i to j (transferring value v) is *possible* and $\langle i, j \rangle$ claims that a communication from i to j is *impossible*.

Note that a no-match claim $\langle i, j \rangle$ implies the *waiting* for a possibility to communicate from i to j. The constraint of no unrealistic waiting that the maximal parallelism model imposes on parallel execution, can now be formulated as: two processes may not make the same no-match claim, i.e., waiting at both sides for the same communication between each other is prohibited.

The communication claim record is the same as the communication record of Francez, Lehmann, and Pnueli (1984). Internal moves within a process (the δ-record of *ibid.*) are modelled by empty bags. The no-match claim record is new and allows

— the checking of the maximal parallelism constraints, i.e., no unnecessary waiting (see above);

— the detection of (established) deadlock (i.e., waiting for a communication that will never come), rendering expectation states as in *ibid.* unnecessary.

Finally, Section 11 contains conclusions and outlines some of the research going on.

2. MINI CSP-R

In this section we describe our language Mini CSP-R. Mini CSP-R consists of the programming constructs of our interest in their basic form without syntactic sugar. In Appendix A we show how Mini CSP-R can easily be extended to a language CSP-R that can simulate the basic Ada real-time and communication primitives.

Mini CSP-R essentially is CSP (see Hoare, 1978) with the addition of the real-time construct **wait** *d*. This construct can be used both as instruction and as guard in a selection or loop. As guard it functions as a time-out, revoking the willingness of a process to communicate (through one of the I/O guards).

In the syntax we use the conventions:

— a *process identification* is an element of $\{P_1, P_2, \ldots\}$,

— a *duration* is an integer-valued expression.

We assume that expressions e and boolean expressions b have some unspecified syntax.

The primitive language elements are the *instructions*, notation Instr:

1. $x := e$ — assignment
2. **wait** d — wait instruction (d is a duration)

3.1 $P_i ! e$ — output (send) to process i the value of the expression e

3.2 $P_i ? x$ — input (receive) from process i a value and assign this value to the variable x.

Instructions of form 3 are called I/O commands: $P_i ! e$ is an output command and $P_i ? x$ an input command.

The important notion of *syntactic matching* of two I/O commands in two processes is defined as follows: two pairs $\langle P_i, \alpha \rangle$ and $\langle P_j, \beta \rangle$ (α, β I/O commands) match syntactically iff ($\equiv$ stands for syntactical equality): ($\alpha \equiv P_j ! e$ and $\beta \equiv P_i ? x$) or ($\alpha \equiv P_j ? x$ and $\beta \equiv P_i ! e$).

Communication between processes i and j takes place when $\langle i, \alpha \rangle$ *semantically matches* $\langle j, \beta \rangle$:

— $\langle P_i, \alpha \rangle$ and $\langle P_j, \beta \rangle$ match syntactically,

— control in P_i and P_j is in front of α, respectively β.

The result of a semantic match is the simultaneous execution of the I/O commands as indicated by 3.1 and 3.2. Its effect is the assignment of the value of the expression of the sending process to the variable of the receiving process.

A *guard* is of one of the forms:

1. b — pure boolean guard
2.1. α — pure I/O guard
2.2. $b; \alpha$ — boolean I/O guard
3.1. **wait** d — pure wait guard
3.2. b; **wait** d — boolean wait guard.

In these clauses, b is a boolean expression (e.g., $x > 0$), α is an I/O command, and d is a duration. For a guard g, its *boolean part* $\bar{g}$ is defined as: $\overline{b} = b$, $\overline{\alpha} = \mathbf{true}$, $\overline{b; \alpha} = b$, $\overline{\mathbf{wait}\ d} = \mathbf{true}$, $\overline{b; \mathbf{wait}\ d} = b$. A guard g is called *open* if $\bar{g}$ evaluates to true.

To complete the definition of Mini CSP-R, we define *commands*, notation Comm, together with *parallel commands*, notation ParComm, and the set of *visible subprocesses* of a command, notation vsp, inductively as follows:

1. every instruction is a command; $\mathrm{vsp}(T) = \varnothing$ for every $T \in \mathrm{Instr}$;

2. if $T_1, T_2 \in \mathrm{Comm}$, then $T_1; T_2$ is a (sequential composition) command with

$$\mathrm{vsp}(T_1; T_2) = \mathrm{vsp}(T_1) \cup \mathrm{vsp}(T_2);$$

3. if $T_1, \ldots, T_n \in \mathrm{Comm}$ and $g_1, \ldots, g_n$ are guards $(n \geqslant 1)$, then $[\Box_{j=1}^{n}\, g_j \to T_j]$ is an (alternative) command and $*\,[\Box_{j=1}^{n}\, g_j \to T_j]$ is a (repetitive) command with

$$\mathrm{vsp}\left(\left[\Box_{j=1}^{n}\, g_j \to T_j\right]\right) = \mathrm{vsp}\left(*\left[\Box_{j=1}^{n}\, g_j \to T_j\right]\right) = \bigcup_{j=1}^{n} \mathrm{vsp}(T_j);$$

4.1. if $T \in \mathrm{Comm}$ and $i > 0$, then $P_i :: T$ is a (named) parallel command;

4.2. if $T_1, T_2 \in \mathrm{ParComm}$ and the following two restrictions are satisfied:

(r1) the variables occurring in T_1 are different from those occurring in T_2,

(r2) the visible subprocesses of T_1 are different from those of T_2,

then $(T_1 \parallel T_2)$ is a (composite) parallel command;

5. a parallel command is also a command with

$$\mathrm{vsp}(P_i :: T) = \{i\} \qquad \text{and} \qquad \mathrm{vsp}((T_1 \parallel T_2)) = \mathrm{vsp}(T_1) \cup \mathrm{vsp}(T_2).$$

Note that in a composite parallel command $(T_1 \| T_2)$ all non-composite commands are of the form $P_i :: T$. We further adopt the naming conventions of (Hoare, 1978; Francez, Lehmann, and Pnueli, 1984): an I/O command within a (named) command $P_i :: T$ may address only one of P_i's sibling processes or one of its ancestor's sibling processes. Note that such a naming convention may result in a match with a subprocess of the named sibling (see Example 5 in Section 8).

We can interpret Mini CSP-R informally as follows (this interpretation applies also to CSP-R):

1.1. An assignment has its usual interpretation: the value of the expression e is assigned to the variable x.

1.2. The wait instruction suspends execution of the process in which it occurs for the value of d (but at least one) time units.

1.3. The interpretation of I/O commands was already indicated above: an I/O command α in process i waits for a semantic match with an I/O command β in a process j.

2. The interpretation of sequential composition is as usual: the execution of T_1 is followed by the execution of T_2.

3.1. The interpretation of an alternative command is as follows: First check if none of the guards is open. If this is the case, execution aborts. Otherwise, check whether there is at least one open pure boolean guard. If this is the case select non-deterministically one of these guards. In the case that at least one of the guards is open but there are no open pure boolean guards, execution of an alternative command proceeds as follows. The waitvalue is defined to be infinite if there are no open wait guards and, otherwise, the maximum of 1 and the minimum of the values of the durations of the open wait guards. For waitvalue time units wait for a semantic match with one of the open I/O guards. As soon as a semantic match occurs within this time period, take it (if more semantic matches occur at the same moment, non-deterministically choose one of them). If no semantic match occurs within waitvalue time units, after this time period one of the open wait guards with a minimal duration is selected. A selection of a guard g_j in all these cases is followed by the execution of the corresponding command T_j.

Observe that in this interpretation of an alternative command a choice has been made; viz., commands guarded by open boolean guards have priority over commands guarded by open I/O guards for which an immediate semantic match is available. This choice is motivated by our aim to model Ada's real-time features (see Appendix A2).

3.2. The interpretation of a repetitive command is the repeated execution of the alternative command contained in it. Now, however,

execution terminates normally whenever in this repetition none of the guards is open.

4.1. The interpretation of a named parallel command is as follows: $P_i :: T$ executes its body T. Furthermore, for a semantic match of *any* I/O command α in T with an I/O command outside T, α is considered to be part of process i and process i only. Hence if α occurs in the body of some visible subprocess of T, α is not addressable by the name of that visible subprocess from outside T anymore. Even more, the visible subprocesses of T are no longer visible outside $P_i :: T$.

4.2. The interpretation of a composite parallel command involves the parallel execution of the parts T_1 and T_2. The underlying parallel execution semantics is *not* interleaving semantics, but a semantics based on the maximal parallelism model (see Sections 3 and 9). For Mini CSP-R this means that whenever there is a choice between different semantic matches for some I/O command in a process, always one of the semantic matches that occurred earliest in time is non-deterministically chosen.

3. The Maximal Parallelism Model

Under maximal parallelism, the number of instructions in concurrently executing processes that can be executed simultaneously without violating synchronization requirements, is maximalized (see Salwicki and Müldner, 1981, for a formal definition). So, for the program $[x := 1 \| x := 3 \| y := 2]$ in some shared variable language *either* the first and third component *or* the second and third component will execute their first move simultaneously, but *not* the first and second component; all this, under the assumption that multiple accesses to a single (shared) variable are mutually exclusive.

Implementing maximal parallelism requires separate processors for the various processes. The connection with real-time behaviour is, that when execution speed is a critical factor, separate processors should be available to all processes. For distributed computing, we take maximal parallelism to mean "first-come first-served" (fcfs) in some global time scale (see Section 4).

Consider the Mini CSP-R program $(P_1 :: (P_{11} :: P_2 ! 0 \| (P_{12} :: P_{13} ! 1 \| P_{13} :: P_{12} ? x; P_2 ! x)) \| P_2 :: P_1 ? y; P_1 ? y)$. According to *interleaving* semantics two scenarios are possible:

(1) P_{11} communicates with P_2 while P_{12} communicates with P_{13}; after that P_{13} communicates with P_2;

(2) P_{12} first communicates with P_{13}; after that P_{13} communicates with P_2; finally, P_{11} communicates with P_2.

According to *maximal parallelism* semantics, only (1) is possible since P_{11} and P_2 can *immediately* become engaged in a rendezvous and hence do not wait for P_{12} and P_{13} to communicate earlier.

The model is, however, not intended to maximize the amount of ongoing activity in a global way. What a process does is decided locally, partially based on the process' knowledge of communications that are being offered to it but otherwise independent of what goes on elsewhere. What the model does guarantee is that whenever a process wants to communicate it will do so at the earliest opportunity and that local noncommunicating actions are executed without any delay.

As we shall reason in Section 9, the maximal parallelism model has some unrealistic aspects for distributed systems in general. We shall develop a whole family of real-time models that range from interleaving to maximal parallelism semantics and that incorporate the transmission time for messages in a system.

4. Our View of Time

To express real-time properties such as "the system responds to a certain request within a fixed number of seconds" there must be some measure of time to relate these properties to. When we talk about abstract, i.e., implementation independent, properties of a system *as a whole*, this measure must be relative to some *global* time scale. For distributed systems this means that all events in the various processes are related to each other by means of one *conceptual* global clock, introduced at a metalevel of reasoning.

Clearly, no physical realization of such a global clock is possible; processors always drift from one time mutual synchronization as exemplified by the existence of clock synchronization algorithms. In our model, drifting can always be modelled by allowing (small) unpredictable variations in the execution time of basic actions.

5. Notations and Technical Preliminaries

This section is intended as a reference to our notation.

5.1. *Numbers, Sets, Cartesian Product and Finite Sequences*

$\mathbb{N} = \{0, 1, 2, \ldots\}$ is the set of natural numbers ordered by $0 < 1 < 2 < \cdots$

$\mathbb{N}^{\infty} = \mathbb{N} \cup \{\infty\}$, inherits the ordering on $\mathbb{N}$ and is additionally ordered by $n < \infty$ for all $n \in \mathbb{N}$.

The empty set is denoted by $\varnothing$.
The powerset of a base set E, i.e., the set of all subsets of E, is denoted by $\mathbf{P}(E)$.
If $E_1, ..., E_n$ are sets, then $X_{i=1}^{n} E_i$ denotes their cartesian product.
If all E_i's are equal (to E), we write E^n for $X_{i=1}^{n} E_i$.
π_i, for $1 \leqslant i \leqslant n$, denote the associated projection functions for elements of $X_{i=1}^{n} E_i$: $\pi_i(\langle e_1, ..., e_n \rangle) = e_i$.
A finite sequence over a base set E is an element of $\mathbf{S}(E) \stackrel{\text{def}}{=} \bigcup_{n \in \mathbb{N}} E^n$, denoted by $\langle e_1, ..., e_n \rangle$ or $\langle e_i \rangle_{i=1}^{n}$, where $e_i \in E$, $1 \leqslant i \leqslant n$.
If all e_i's are equal (to e), we write $\langle e \rangle^n$ for $\langle e_i \rangle_{i=1}^{n}$.
A special case is $n = 0$: it is called the empty sequence, notation λ.
The length of a sequence $s = \langle e_1, ..., e_n \rangle$, notation $|s|$, is n.
For a sequence $s = \langle e_1, ..., e_n \rangle$ and $1 \leqslant k \leqslant n$ we define the kth element of s, notation $s(k)$, as e_k.
For $e \in E$ and $s \in \mathbf{S}(E)$, we say that e is an element of s, notation $e \lessdot s$, if there exists a k, $1 \leqslant k \leqslant |s|$, such that $s(k) = e$.
Given $s_1, s_2 \in \mathbf{S}(E)$, we can concatenate them, notation $s_1 \hat{} s_2$: if $s_1 = \langle e_1, ..., e_n \rangle$ and $s_2 = \langle e'_1, ..., e'_m \rangle$, then $s_1 \hat{} s_2 = \langle e_1, ..., e_n, e'_1, ..., e'_m \rangle$. Note that $\hat{}$ is closed, associative, and has identity element λ:

$$s_1, s_2 \in \mathbf{S}(E) \Rightarrow s_1 \hat{} s_2 \in \mathbf{S}(E), \qquad (s_1 \hat{} s_2) \hat{} s_3 = s_1 \hat{} (s_2 \hat{} s_3),$$

and

$$s \hat{} \lambda = \lambda \hat{} s = s.$$

For $s, s' \in \mathbf{S}(E)$ we say that s' is a prefix of s, notation $s' \leqslant s$, if there exists a $s'' \in \mathbf{S}(E)$ such that $s = s' \hat{} s''$.

5.2. *Functions and Partial Functions*

The set of all functions from X (the domain) to Y (the range) is denoted by Y^X. The domain and range of a function f are denoted by $\mathrm{dom}(f)$ (resp. $\mathrm{ran}(f)$). A partial function from X to Y is an element of $Y^{X'}$, where $X' \in \mathbf{P}(X)$, i.e., a function from a subset of X to Y.

For f a partial function from X to Y, $x \in X$ and $y \in Y$, $f[y/x]$ is the partial function with $\mathrm{dom}(f[y/x]) = \mathrm{dom}(f) \cup \{x\}$ and $\mathrm{ran}(f[y/x]) = Y$ defined by

$$(f[y/x])(x') = \begin{cases} y & \text{if } x' = x, \\ f(x') & \text{if } x' \in \mathrm{dom}(f) \setminus \{x\}. \end{cases}$$

5.3. *Bags.*

A bag (or multiset) over a base set E is an element of $\mathbf{B}(E) \stackrel{\text{def}}{=} \mathbb{N}^E$, i.e., a function from E to $\mathbb{N}$. For $e \in E$ and $B \in \mathbf{B}(E)$ we say that e is an element of B, notation $e \in B$, if $B(e) > 0$.

For finite bags we often use the notation $[e_1^{i_1}, ..., e_n^{i_n}]$, where $n \in \mathbb{N}$, $i_k \geqslant 1$, $e_k \in E$, all e_k different $(1 \leqslant k \leqslant n)$ which corresponds to the bag $B \in \mathbf{B}(E)$ defined by

$$B(e) = \begin{cases} i_k & \text{if } e = e_k, 1 \leqslant k \leqslant n, \\ 0 & \text{otherwise.} \end{cases}$$

If $i_k = 1$, we just write e_k instead of e_k^1. A special case is $n = 0$, the empty bag, notation [ ].

6. The Semantic Domain and Its Interpretation

6.1. *The Semantic Domain*

Because our basic domain consists of state-history pairs, we first explain what states and histories are:

Let Id be a (fixed) set of identifiers (i.e., a set of strings over some alphabet). Since we gave no syntax for expressions in Mini CSP-R, we assume furthermore the existence of a set V of expression values.

S, the set of *proper* states, is defined to be the set of partial functions from Id to V. So a proper state $s \in S$ maps *certain* identifiers to their value.

Σ, the total set of states, can now be defined as $S \cup \{\perp, \bullet\}$, where $\perp$ denotes an incomplete computation and $\bullet$ denotes failure (both explained later).

Let $\mathrm{CAR} = (\mathbb{N} \times \mathbb{N}) \cup (\mathbb{N} \times \mathbb{N} \times V)$ be the set of communication assumption records (for the intuition, see the last part of the Introduction).

H, the set of histories, is, as was motivated in the Introduction, $\mathbf{S}(\mathbf{B}(\mathrm{CAR}))$. It would in fact suffice to take $H = \mathbf{S}(\mathbf{P}(\mathbb{N} \times \mathbb{N}) \times \mathbf{B}(\mathbb{N} \times \mathbb{N} \times V))$, as bags are only needed to collect communication claims. Obviously, for claiming the *absence* of a communication possibility between processes i and j, it suffices to do this only once. However, we prefer the first notationally simpler definition. The technical reason for using bags instead of sets is illustrated in Example 3 of Section 8. Our central domain is that of non-empty prefix-closed sets of state-history pairs, notation ΣH:

Definition. A set $X \in \mathbf{P}(\Sigma \times H)$ is *prefix-closed* iff for all $\langle \sigma, h \rangle \in X$, if $h' \leqslant h$, then $\langle \perp, h' \rangle \in X$. Define $\Sigma H = \{X \in \mathbf{P}(\Sigma \times H) \mid X \neq \varnothing$ and X is prefix-closed$\}$. The *prefix-closure* of X, $\mathrm{PFC}(X)$, is defined as

$$X \cup \{\langle \perp, \lambda \rangle\} \cup \{\langle \perp, h' \rangle \mid \exists \sigma \, \exists h (\langle \sigma, h \rangle \in X \wedge h' \leqslant h)\}.$$

Note that $\mathrm{PFC}(X) \in \Sigma H$, for all $X \in \mathbf{P}(\Sigma \times H)$. ΣH can be turned into a complete lattice:

— the partial ordering is $\subseteq$, set-inclusion;

— the least upper bound is obtained by $\cup$, set-union.

Its least element is $\{\langle \perp, \lambda \rangle\}$.

The technical motivation for the introduction of $\perp$ lies in the simplicity of the ordering of ΣH: several proofs, in particular those for continuity of operators, become very simple.

The introduction of a separate failure state $\bullet$ is needed for the detection of non-deterministic failure (see below, in Section 6.2).

We want the elements of ΣH to be non-empty, because otherwise the least element of ΣH would be $\varnothing$. Since $\varnothing$ contains no history at all, and sequential composition is essentially modelled by concatenation of histories, this choice of least element would imply that the denotation of $*[\mathbf{true} \to P_2!\,5]$ would be empty. Although consistent with the view that a command is a transformation of initial states to final states when characterizing sequential constructs relationally, this does not capture our intuition that an unbounded set of communication possibilities may have been offered by $*[\mathbf{true} \to P_2!\,5]$ (cf. Example 1 in Section 8).

Remark. As E.-R. Olderog observed in the context of the linear history semantics for CSP (see Francez, Lehmann, and Pnueli, 1984), here too, we do not need to order our domain. This is a consequence of the fact that our recursions are always guarded (see loops) and that histories, once they have been generated, can not "shrink," i.e., they remain the same or are extended to a longer history. For details, see the Appendix of *ibid.*

6.2. *Interpretation of* ΣH

We can interpret $X \in \Sigma H$ as the set of all possible computations of a program P (cf. Francez, Lehmann, and Pnueli, 1984):

— $\langle s, h \rangle \in X$ with $s \in S$, models a computation of P producing history h that terminates in s;

— $\langle \bullet, h \rangle \in X$ models a failure of P after producing history h;

— $\langle \perp, h \rangle \in X$ models an incomplete computation of P which is *either* an approximation of a computation $\langle \sigma, h' \rangle$ with $\sigma \neq \perp$ and $h \leqslant h'$ *or* an element in a chain of approximations $\langle \perp, h_0 \rangle, \langle \perp, h_0 \hat{} h_1 \rangle, \cdots$ (all $h_i \neq \lambda$) which models an *infinite* computation of P with history $h_0 \hat{} h_1 \hat{} \cdots$ (this interpretation can be justified by an appeal to König's lemma, based on an intuitive operational semantics).

If only deterministic failures can occur, there is no need for a separate

failure state • because $\perp$ can be used for that purpose: deterministic failure of P after history h is then modelled by $\langle \perp, h \rangle \in X$ such that there exists neither $\langle s, h' \rangle \in X$ with $s \in S$, $h \leqslant h'$ nor $\langle \perp, h' \rangle \in X$ with $h \leqslant h'$, $h \neq h'$. However, we have to include the possibility of non-deterministic failure as demonstrated by the following Mini CSP-R program fragment: $[\mathbf{true} \rightarrow [\mathbf{false} \rightarrow x := 0] \Box \mathbf{true} \rightarrow x := 1]$.

Using the above interpretation of ΣH, we can informally define a notion of observable behaviour. The observable entities are: a communication history, termination, failure, and infinite computation. The observable behaviour of a communication history has already been given in the Introduction. The other observable entities are given in the above interpretation of ΣH:

— termination: indicated by a proper state $s \in S$,

— failure: indicated by •,

— infinite computation: indicated by an infinite chain of approximations.

Both divergence and established deadlock are viewed as infinite computations: divergence is making internal steps while time passes, established deadlock is waiting for a communication that will not come, while time passes. This means that divergence and established internalized deadlock are observed in the same way, and hence cannot be distinguished. In our view this is a perfectly reasonable standpoint: the only observation that can be made from the outside is the ticking of the global clock while no communication with the environment can occur. In other words: there is no context that can distinguish a diverging process from such a deadlocked one (cf. Example 2 in Section 8).

7. Maximal Parallelism Semantics for Mini CSP-R

7.1. *Introduction*

The meaning of Mini CSP-R commands is defined *denotationally* by giving for all commands T, an equation which relates the meaning of T, notation $\mathbf{M}[\![T]\!]$, to the meaning of T's constituents in a compositional way. In Section 7.2 we show that it suffices to define $\mathbf{M}[\![T]\!]$ as a function from S to ΣH.

To define the alternative command $[\Box_{j=1}^{n} g_j \rightarrow T_j]$ compositionally, we use an auxiliary semantic function $\mathbf{G}[\![g, A]\!]$ from S to ΣH which gives the meaning of guard g in the context of a set A of alternative guards (the other guards in the alternative command). We use the context A in a compositional way, i.e., A depends only on the alternative command in which g

occurs. **G** is furthermore used in defining the meaning of guards that occur as instructions (these are the pure waitguards and pure I/O guards). The meaning of such an instruction is simply the meaning of the guard in an empty context.

Since we gave no syntax for (boolean) expressions in Mini CSP-R, we assume the existence of semantic functions **V** and **W**, such that $\mathbf{V}[\![e]\!]$ for e an expression is a function from S to V, and $\mathbf{W}[\![b]\!]$ for b a boolean expression is a predicate on S, i.e., for $s \in S$, $\mathbf{W}[\![b]\!]s$ is either true or false.

To define the meaning of constructs like $P_1 :: P_2!5$ compositionally, we have to give a meaning for $P_2!5$ separately, i.e., in a context where it is not known that this construct belongs to the process with identification 1. In order to do so, we introduce as semantic entity the "unknown process," with process identification 0, and use this, e.g., to generate records $\langle 0, 2, 5 \rangle$ in the meaning for $P_2!5$ and later, in the meaning for $P_1 :: P_2!5$, replace 0 by 1. Therefore, we identify process identifications with natural numbers.

Just as for the syntax we need a notion of visible subprocesses of a command T, $\mathrm{VS}(T)$. The difference with the definition in Section 2 is the use of $\{0\}$ instead of $\varnothing$:

$$\mathrm{VS}(T) = \{0\} \text{ for } T \text{ an instruction},$$

$$\mathrm{VS}(T_1; T_2) = \mathrm{VS}((T_1 \parallel T_2)) = \mathrm{VS}(T_1) \cup \mathrm{VS}(T_2),$$

$$\mathrm{VS}\left(\left[\Box_{j=1}^{n} g_j \to T_j\right]\right) = \mathrm{VS}\left(*\left[\Box_{j=1}^{n} g_j \to T_j\right]\right) = \{0\} \cup \bigcup_{j=1}^{n} \mathrm{VS}(T_j),$$

$$\mathrm{VS}(P_i :: T) = \{i\}.$$

In the third line, the zero is needed to account for I/O guards as, e.g., in $P_1 :: [P_2!0 \to P_3 :: x := 0]$.

To keep the semantics simple, we assume that the evaluation of expressions takes no time. However, this restriction can easily be relaxed by introducing time-parameters that represent evaluation times of expressions. Furthermore, we make the realistic assumption that the execution of commands takes at least one unit of time unless failure occurs (this can only occur if an alternative command which has no open guard is executed). The idea behind this decision is that we want to exclude the unrealistic possibility of an infinite loop taking zero time. Such a loop is possible in Ada, as shown in Appendix A2, and obviously this possibility must be excluded. Appendix A2 contains a discussion how to do so.

7.2. *Extending the Meaning Function*

$\mathbf{M}[\![T]\!]$, the meaning of a construct T, only depends on a proper state $s \in S$: $\mathbf{M}[\![T]\!]s \in \Sigma H$ represents all possible state changes and computational

histories produced by T starting from s. It therefore seems sufficient to let $\mathbf{M}[\![T]\!]$ be a function from S to ΣH. However, to define sequential composition we have to extend the meaning function to a function from ΣH to ΣH (this situation is analogous to that for a purely sequential non-deterministic language where the meaning function is generalized to sets of states). This extension shall be defined uniformly for all functions from S to ΣH, so we can still use $\mathbf{M}[\![T]\!]$ as a function from S to ΣH keeping in mind that this extension must be used when composing meaning functions. We first extend a function ϕ from S to ΣH to a function ϕ^+ from Σ to ΣH and next to a function ϕ^* from ΣH to ΣH.

DEFINITION. Let ϕ be a function from S to ΣH. Then ϕ^+ is the function from Σ to ΣH defined by

$$\phi^+(\sigma) = \begin{cases} \phi(\sigma), & \text{if } \sigma \in S, \\ \mathrm{PFC}(\{\langle \sigma, \lambda \rangle\}), & \text{otherwise.} \end{cases}$$

Furthermore, ϕ^* is the function from ΣH to ΣH defined by

$$\phi^*(X) = \{\langle \sigma', h \hat{} h' \rangle \mid \langle \sigma, h \rangle \in X \wedge \langle \sigma', h' \rangle \in \phi^+(\sigma)\}.$$

ϕ^* extends ϕ in a canonical way: for $X \in \Sigma H$ it takes $\langle \sigma, h \rangle \in X$ and extends h with an additional history h' formed by applying ϕ^+ to σ; ϕ^+ behaves like ϕ on S but takes care that histories of pairs $\langle \sigma, h \rangle \in X$ with $\sigma \notin S$ are not extended; the new state σ' is the state after applying ϕ^+ to σ.

The histories h represent communication assumptions that have been made and can only be supplemented with additional communication assumptions. In other words: the extension of histories is independent of their contents. The meaning function should certainly have this property. A property of ϕ^* is that it is always strict and continuous, as proved below. This means that we do not have to worry about the continuity of operators in our semantics!

PROPOSITION. *For all ϕ from S to ΣH, ϕ^* is a strict and continuous function from ΣH to ΣH.*

Proof. $\phi^*(\{\langle \perp, \lambda \rangle\}) = \{\langle \sigma', \lambda \hat{} h' \rangle \mid \langle \sigma', h' \rangle \in \phi^+(\perp)\} = \phi^+(\perp) = \{\langle \perp, \lambda \rangle\}$ and

$$\begin{aligned} \phi^*\left(\bigcup_{i \in I} X_i\right) &= \left\{\langle \sigma', h \hat{} h' \rangle \mid \langle \sigma, h \rangle \in \bigcup_{i \in I} X_i \wedge \langle \sigma', h' \rangle \in \phi^+(\sigma)\right\} \\ &= \bigcup_{i \in I} \{\langle \sigma', h \hat{} h' \rangle \mid \langle \sigma, h \rangle \in X_i \wedge \langle \sigma', h' \rangle \in \phi^+(\sigma)\} \\ &= \bigcup_{i \in I} \phi^*(X_i). \quad \blacksquare \end{aligned}$$

7.3. *Definition of* **G**

In the definition of **G** we use the following two auxiliary notions for guards:

DEFINITION 1. For a set of guards G and $s \in S$, define $\mathrm{RTA}(G, s) \in \mathbf{B}(\mathrm{CAR})$, the bag of real-time assumptions concerning the open I/O guards of G in state s as follows:

$$\mathrm{RTA}(G, s)(r) = \begin{cases} 1 & \text{if } r \in \{\langle 0, i\rangle \mid \exists g \in G(g \equiv P_i!\, e \vee (g \equiv b; P_i!\, e \wedge \mathbf{W}[\![b]\!]s))\} \\ & \quad \cup \{\langle i, 0\rangle \mid \exists g \in G(g \equiv P_i?\, x \vee (g \equiv b; P_i?\, x \wedge \mathbf{W}[\![b]\!]s))\}, \\ 0 & \text{otherwise.} \end{cases}$$

Remark. If, e.g., $P_2!\,4$ and $P_2!\,6$ occur in G one might expect a multiplicity 2 (instead of 1) for the record $\langle 0, 2\rangle$ in the above definition. This is unnecessary (see the discussion of bags versus sets in Section 6.1).

DEFINITION 2. For a guard g and $s \in S$, define $\mathrm{waitvalue}(g, s) \in \mathbb{N}^\infty$ as follows:

$$\mathrm{waitvalue}(g, s) = \begin{cases} 0 & \text{if } g \equiv b \wedge \mathbf{W}[\![b]\!]s, \\ \max\{\mathbf{V}[\![d]\!]s, 1\} & \text{if } g \equiv \mathbf{wait}\ d \vee (g \equiv b; \mathbf{wait}\ d \wedge \mathbf{W}[\![b]\!]s), \\ \infty & \text{otherwise.} \end{cases}$$

Furthermore, for a set of guards G and $s \in S$, define $\mathrm{minwait}(G, s) \in \mathbb{N}^\infty$ as $\min\{\mathrm{waitvalue}(g, s) \mid g \in G\}$ (where by convention $\min \varnothing = \infty$).

Note that the guard **true** has waitvalue 0 while the guards **wait** 0 and **wait** 1 have waitvalue 1. The decision to let **wait** 0 have waitvalue 1 is explained in Appendix A2.

The equations for **G** are (see Section 7.1 for its use and motivation):

$$\mathbf{G}[\![b, A]\!]s = \begin{cases} \mathrm{PFC}(\{\langle s, \lambda\rangle\}) & \text{if } \mathbf{W}[\![b]\!]s, \\ \{\langle \bot, \lambda\rangle\} & \text{otherwise.} \end{cases}$$

A boolean acts as a filter: s is maintained only if b evaluates to true in s.

$$\mathbf{G}[\![\mathbf{wait}\ d, A]\!]s = \mathrm{PFC}(\{\langle s, \langle \mathrm{RTA}(A, s)\rangle^T\rangle \mid \max\{\mathbf{V}[\![d]\!]s, 1\} = \mathrm{minwait}(A \cup \{\mathbf{wait}\ d\}, s) \stackrel{\mathrm{def}}{=} T\}).$$

A pure wait guard in the context A can be selected after its waitvalue time units elapsed provided this value equals the minimal waitvalue T

(note that $T \in \mathbb{N}$) *and* no semantic match for an open I/O guard in A occurred in this period. If there is at least one open boolean guard in A, then $T = 0$ and no wait guard can be selected.

$$\mathbf{G}[\![P_j!\, e, A]\!]s = \mathrm{PFC}(\{\langle s, \langle \mathrm{RTA}(\mathrm{GRDS}, s)\rangle^t \hat{}\ \langle[\langle 0, j, \mathbf{V}[\![e]\!]s\rangle]\rangle \mid 0 \leqslant t < \mathrm{minwait}(A, s)\}),$$

where $\mathrm{GRDS} = A \cup \{P_j!\, e\}$.

A pure I/O guard in the context A can be selected (indicated by the last triple of the history above) within the minimum waitvalue of A (the bound on t above) under the condition that no semantic match for any open I/O guard in GRDS occurred earlier (indicated by the first t elements of the history above). If there is at least one open boolean guard in A, then $\mathrm{minwait}(A, s) = 0$ and no output guard (in fact, no I/O guard) can be selected. The possibility that no guard at all is selected can only occur if there are no open boolean guards and no open wait guards (hence $\mathrm{minwait}(A, s) = \infty$) and furthermore no semantic match for an open I/O guard ever occurs. This case is represented by the subset $\{\langle \bot, \langle \mathrm{RTA}(\mathrm{GRDS}, s)\rangle^t\rangle \mid t \in \mathbb{N}\}$ of $\mathbf{G}[\![P_j!\, e, A]\!]s$ (remember, this is a prefix-closed set).

$$\mathbf{G}[\![P_j?\, x, A]\!]s = \mathrm{PFC}(\{\langle s[v/x], \langle \mathrm{RTA}(\mathrm{GRDS}, s)\rangle^t \hat{}\ \langle[\langle j, 0, v\rangle]\rangle \mid v \in V, 0 \leqslant t < \mathrm{minwait}(A, s)\}),$$

where $\mathrm{GRDS} = A \cup \{P_j?x\}$.

The same remarks as for $\mathbf{G}[\![P_j!\, e, A]\!]s$ apply here. In comparison with $\mathbf{G}[\![P_j!\, e, A]\!]s$ we see that in the last triple of the history sender and receiver are reversed. Furthermore, for an input command $P_j?\, x$ we have to "guess" the value v that will be assigned to x. When binding the inputting process with the outputting process we check that the values correspond (see the last three examples in Section 8). This "guessing" models Bekiç' and Milner's concept of renewal (see Milner, 1973).

$$\mathbf{G}[\![b; g, A]\!]s = \mathbf{G}[\![g, A]\!]^* (\mathbf{G}[\![b, A]\!]s), \qquad \text{where } g \equiv P_j!\, e \text{ or } g \equiv P_j?\, x \text{ or } g \equiv \mathbf{wait}\ d.$$

The meaning of a sequential composition of guards is the functional composition (using the extension operator "*") of the meaning of the separate guards.

7.4. *Definition of* **M**

7.4.1. $\mathbf{M}[\![T]\!]$ *for* $T \in \mathrm{Comm} \backslash \mathrm{ParComm}$. In this subsection we give the meaning of the non-parallel commands of Mini CSP-R.

$$\mathbf{M}[\![x := e]\!]s = \mathrm{PFC}(\{\langle s[\mathbf{V}[\![e]\!]s/x], \langle [\ ] \rangle \rangle\}).$$

To keep the semantics simple, an assignment takes exactly one time unit (indicated by the empty bag).

$$\mathbf{M}[\![g]\!]s = \mathbf{G}[\![g, \varnothing]\!]s, \qquad \text{for} \quad g \equiv \textbf{wait}\ d \text{ or } g \equiv P_j!\, e \text{ or } g \equiv P_j?\, x.$$

This use of **G** was already discussed in Section 7.1.

$$\mathbf{M}[\![T_1; T_2]\!]s = \mathbf{M}[\![T_2]\!]^* (\mathbf{M}[\![T_1]\!]s).$$

$$\mathbf{M}\left[\!\!\left[\left[\ \Box_{j=1}^{n}\ g_j \to T_j\right]\right]\!\!\right]s$$

$$= \begin{cases} \bigcup_{j=1}^{n} \mathbf{M}[\![T_j]\!]^* (\mathbf{G}[\![g_j, \{g_k \mid 1 \leqslant k \leqslant n,\ k \neq j\}]\!]s) & \text{if } \bigvee_{j=1}^{n} \mathbf{W}[\![\bar{g}_j]\!]s, \\ \mathrm{PFC}(\{\langle \bullet, \lambda \rangle\}) & \text{otherwise.} \end{cases}$$

The meaning of the alternative command depends on the presence of an open guard: if no such guard is present this means failure, otherwise one guard is selected where each guard is considered in the context of the remaining guards ($\bar{g}_j$ is the boolean part of g_j, see Section 2).

Let C abbreviate $[\Box_{j=1}^{n}\ g_j \to T_j]$.

$$\mathbf{M}[\![* \mathrm{C}]\!]s = \bigcup_{i \in \mathbb{N}} \phi_i(s),$$

where the ϕ_i ($i \in \mathbb{N}$) are functions from S to ΣH defined inductively by

$$\phi_0(s) = \{\langle \bot, \lambda \rangle\} \qquad \text{for all} \quad s \in S,$$

$$\phi_{i+1}(s) = \begin{cases} \phi_i^*(\mathbf{M}[\![C]\!]s) & \text{if } \bigvee_{j=1}^{n} \mathbf{W}[\![\bar{g}_j]\!]s, \\ \mathrm{PFC}(\{\langle s, \langle [\ ] \rangle \rangle\}) & \text{otherwise.} \end{cases}$$

The ϕ_i's represent as usual the ith iteration step of the loopbody. If at some point of iteration there are no open guards anymore, the loop terminates (this last iteration is indicated by the empty bag because the execution of commands takes at least one time unit).

For an illustration of the loop equation see the first two examples in

Section 8 (these give also a demonstration why $\{\langle \perp, \lambda \rangle\}$ and not $\varnothing$ should be the least element of ΣH). The loop equation can alternatively be written as a fixed-point equation over the complete partial order of functions from S to ΣH with the usual ordering on function domains:

$$\mathbf{M}[\![* \, C]\!] = \mu(\lambda\phi.\lambda s. \textbf{ if } \bigvee_{j=1}^{n} \mathbf{W}[\![\bar{g}_j]\!] s \textbf{ then } \phi^*(\mathbf{M}[\![C]\!] s) \textbf{ else } \mathrm{PFC}(\{\langle s, \langle [\ ] \rangle \rangle\}) \textbf{ fi}),$$

where μ is the least fixed-point operator.

7.4.2. *The meaning of* $P_i :: T$. The effect caused by $P_i :: T$ is the renaming of the visible subprocesses of T by i. To this end, we need a definition for substitution of a certain process, in this case i, in place of a collection of processes I, in this case $\mathrm{VS}(T)$, both for bags over CAR as for elements of ΣH. Although the substitution for bags over CAR is intuitively clear, the technical definition is rather awkward and is therefore given in Appendix B. So, assuming we have defined $B[I \to i] \in \mathbf{B}(\mathrm{CAR})$ for $B \in \mathbf{B}(\mathrm{CAR})$, $I \in \mathbf{P}(\mathbb{N})$ and $i \in \mathbb{N}$, we can extend this componentwise to elements of ΣH:

$$X[I \to i] = \{\langle \sigma, \langle h(k)[I \to i] \rangle_{k=1}^{|h|} \rangle \mid \langle \sigma, h \rangle \in X\}.$$

LEMMA. *$X[I \to i] \in \Sigma H$ for all $X \in \Sigma H$, $I \in P(\mathbb{N})$ and $i \in \mathbb{N}$.*

Proof. $X[I \to i]$ non-empty: $X \in \Sigma H$ implies $\langle \perp, \lambda \rangle \in X$ and hence $\langle \perp, \lambda \rangle \in X[I \to i]$.

$X[I \to i]$ prefix-closed: Let $\langle \sigma, h \rangle \in X$ and $h' \leqslant \langle h(k)[I \to i] \rangle_{k=1}^{|h|}$. Then $|h'| \leqslant |h|$, so there exists a $h'' \leqslant h$ with $|h''| = |h'|$. Because $X \in \Sigma H$ it follows that $\langle \perp, h'' \rangle \in X$ and hence $\langle \perp, h' \rangle = \langle \perp, \langle h(k)[I \to i] \rangle_{k=1}^{|h'|} \rangle = \langle \perp, \langle h(k)[I \to i] \rangle_{k=1}^{|h''|} \rangle = \langle \perp, \langle h''(k)[I \to i] \rangle_{k=1}^{|h''|} \rangle \in X[I \to i]$. ∎

Now we can define

$$\mathbf{M}[\![P_i :: T]\!] s = (\mathbf{M}[\![T]\!] s)[\mathrm{VS}(T) \to i].$$

7.4.3. *The meaning of* $(T_1 \parallel T_2)$.

7.4.3.1. *Intuition for parallel composition.* It remains to define the meaning of the most important construct, the parallel composition. Intuitively, when binding two processes, the information of the states is combined, the histories are checked for consistency, and then they are merged. Actually this consistency check can be split into two independent parts to be applied at each instant of time:

(c1) Check that histories have matching communication claims, i.e., that histories agree on the communications that occur between the two processes (their internal communications).

(c2) Check that there is no unnecessary waiting, i.e., that histories do not indicate a situation where both processes are waiting for a communication that the other process can provide (in other words: two processes do not wait if there is a semantic match between them).

Check (c1) is the communication consistency check for CSP as in (Francez, Lehmann, and Pnueli, 1984). We call (c2) the real-time consistency check because it enforces maximal parallelism (see the end of Section 1). Since the equation for $\mathbf{M}[\![(T_1 \| T_2)]\!]s$ is rather complex, we give the intuition behind its steps below, and postpone its formal definition till Section 7.4.3.6.

To combine the meanings of $\mathbf{M}[\![T_1]\!]s$ and $\mathbf{M}[\![T_2]\!]s$ to $\mathbf{M}[\![(T_1 \| T_2)]\!]s$, first the states of $\mathbf{M}[\![T_1]\!]s$ and $\mathbf{M}[\![T_2]\!]s$ should be combined. Although trivial at first sight, this raises problems since we can not always assume that such states have disjoint domains, as illustrated by the program $x := 0$; $(P_1 :: x := 1 \| P_2 :: y := 2)$. This is solved in Section 7.4.3.2.

Next consistency checks (c1) and (c2) must be applied to the communication assumption records in $\mathbf{M}[\![T_1]\!]s$ and $\mathbf{M}[\![T_2]\!]s$. Note that for (c1) it is desired to have a *common* communication claim record in both histories while, on the contrary, (c2) checks that there is *no common* no-match claim record in both histories. Moreover, our semantics is such that in the records in the generated histories of a command always at least one of the processes involved is a visible subprocess of that command (see the History Property in Section 7.5). Consequently, for (c2) it is sufficient to check for the absence of identical no-match claims. For (c1), however, one first must establish the visible subprocesses of T_1 and T_2 prior to checking whether a communication claim record in one history should be complemented by an identical record in the other history (since a visible subprocess of T_1 *may* address a process outside of T_2). Therefore, it would be nice if we could first merge the histories that are consistent according to (c2) and *after that* check (c1). Unfortunately this is unfeasible, as is illustrated by the programs $(P_1 :: P_2!0 \| P_2 :: P_1?x)$ and $(P_1 :: (P_{11} :: P_2!0 \| P_{12} :: P_2!0) \| P_2 :: x := 0)$. When following the above approach, the semantics of both these programs would contain the history $\langle [\langle 1, 2, 0 \rangle^2] \rangle$. Now, this history should represent both a successful communication (the first program) *and* deadlock (the second program): an impossibility. We solve this problem through first subtracting equal communication claim records from each other, and after that check whether any internal communication claims are left. Together with the definition

of the real-time consistency check (c2), this is worked out in detail in Section 7.4.3.3.

Third, not all histories should be compared when merging. When combining state-history pairs with $\perp$ as state component(s), representing incomplete computation, special care should be taken to guarantee that indeed *all* the events occuring at a particular time are collected in the resulting history. E.g., $\langle \perp, \lambda \rangle \in \mathbf{M}[\![P_1 :: P_3 ! 5]\!] s$ should not be merged with $\langle s[0/x], \langle [\ ] \rangle \rangle \in \mathbf{M}[\![P_2 :: x := 0]\!] s$, because the result $\langle \perp, \langle [\ ] \rangle \rangle$ will not represent the attempt of P_1 to communicate with P_3 at time 1. This is treated in Section 7.4.3.4.

As the last step, when giving the meaning of $(T_1 \| T_2)$ in terms of its components, the real-time assumptions (represented by the no-match claim records) concerning the visible subprocesses of T_1 and T_2 should be checked and removed. This is illustrated by the program $(P_1 :: P_2 ! 5 \| P_2 :: P_1 ! 5)$. Some histories of P_1 contain the no-match claim $\langle 1, 2 \rangle$, and some of P_2 the no-match claim $\langle 2, 1 \rangle$. After binding P_1 and P_2, the real-time assumptions concerning the collection of processes $\{1, 2\}$ should be checked; in this case, exactly $\langle 1, 2 \rangle$ and $\langle 2, 1 \rangle$. After this check they are not needed anymore and can be removed, since it has been established that no communication will occur.

These four steps correspond with those of the definition of $\mathbf{M}[\![(T_1 \| T_2)]\!] s$, in that order.

7.4.3.2. *Combining states.* For $\mathbf{M}[\![(T_1 \| T_2)]\!] s$, the states of $\mathbf{M}[\![T_1]\!] s$ and $\mathbf{M}[\![T_2]\!] s$ should be combined. Because of the syntactic restriction that the variables of T_1 and T_2 are disjoint (see Section 2, definition of commands), it seems that one can simply form the disjoint union of such states. This is, however, not the case: the state s of the computation up till now can cause problems. For example, in the program $x := 0$; $(P_1 :: x := 1 \| P_2 :: y := 2)$, x is defined *both* in P_1 and P_2. Fortunately, this is only the case for variables that were defined earlier in the program, or in other words: variables that belong to the domain of s. Variables outside the domain of s belong either to P_1 or P_2 (because of the above-mentioned syntactic restriction). The union of states of $\mathbf{M}[\![T_1]\!] s$ and $\mathbf{M}[\![T_2]\!] s$ can now be defined relative to $s \in S$:

Let for $1 \leqslant i \leqslant 2$, $s_i \in S$ belong to $\mathbf{M}[\![T_i]\!] s$ (then $\mathrm{dom}(s) \subseteq \mathrm{dom}(s_i)$). Define the union of s_1 and s_2 relative to s, notation $s_1 \cup_s s_2$, as follows:

$$\mathrm{dom}(s_1 \cup_s s_2) = \mathrm{dom}(s_1) \cup \mathrm{dom}(s_2)$$

and

$$(s_1 \cup_s s_2)(x) \stackrel{\text{def}}{=} s_i(x) \quad \text{if } x \in \mathrm{dom}(s_i) \backslash \mathrm{dom}(s) \text{ or } x \in \mathrm{dom}(s), s(x) = s_{3-i}(x).$$

As remarked above, if $x \in \mathrm{dom}(s_i) \backslash \mathrm{dom}(s)$ then $x \notin \mathrm{dom}(s_{3-i})$. In that case, x is a new variable of T_i and the value of that variable in the combined state is $s_i(x)$. For example, for $t := 0$; $(P_1 :: y := 1 \parallel P_2 :: z := 2)$, $\mathrm{dom}(s) = \{t\}$, $\mathrm{dom}(s_1) = \{t, y\}$, and $\mathrm{dom}(s_2) = \{t, z\}$.

On the other hand, if $x \in \mathrm{dom}(s)$, then at most one of T_1 and T_2 can use x, hence $s_i(x) = s(x)$ for $i = 1$ or $i = 2$. In this case, the value of x in the combined state is $s_{3-i}(x)$. For example, for $t := 0$; $(P_1 :: t := 1 \parallel P_2 :: z = 2)$, $\mathrm{dom}(s) = \{t\}$, $\mathrm{dom}(s_1) = \{t\}$, $\mathrm{dom}(s_2) = \{t, z\}$, and the value of t after this program is 1. Note that $\bigcup_s$ (for all $s \in S$) is commutative and associative.

It remains to extend $\bigcup_s$ for s_i that belong to $\mathbf{M}[\![T_i]\!]s$ but with s_1 or s_2 (or both) not in S. The idea is that whenever one of the s_i represents an incomplete computation the combination represents the same; otherwise, when one of the states represents failure, the combination represents failure:

$$\bot \cup_s \sigma = \sigma \cup_s \bot = \bot \qquad \text{for all} \quad s \in S, \sigma \in \Sigma$$

and

$$\bullet \cup_s \sigma = \sigma \cup_s \bullet = \bullet \qquad \text{for all} \quad s \in S, \sigma \in \Sigma \backslash \{\bot\}.$$

Note that this extension maintains commutativity and associativity.

7.4.3.3. *The consistency check.* There is a direct correspondence between the two parts of the consistency check and the two types of communication assumption records:

(c1) concerns triples $\langle i, j, v \rangle$ such that i and j are internal processes, i.e., processes that belong to the collection of processes represented by the two histories whose consistency is checked; check (c1) corresponds to: each such triple in one history should also occur in the other history at the same time and vice versa

(c2) concerns pairs $\langle i, j \rangle$; it corresponds to: no pair $\langle i, j \rangle$ in one history may occur at the same time in the other history.

Note that for (c1) we need to know the set of internal processes while this is not necessary for (c2). The reason for this is that in all records in the histories generated by our semantics one of the processes i and j refers to the process that generated this record (this history property is proved in Section 7.5). Because (c2) checks that two histories representing different processes do not contain at the same time a common record $\langle i, j \rangle$, this means that i and j must be internal processes anyway.

The real-time consistency check (c2) is formulated by

$$h_1 \not\in^{\mathbf{RT}} h_2 \stackrel{\text{def}}{=} \neg \exists i, j, k \in \mathbb{N}\ (1 \leqslant k \leqslant \min\{|h_1|, |h_2|\} \wedge \langle i, j \rangle \in h_1(k) \wedge \langle i, j \rangle \in h_2(k)).$$

Of course, the consistency check as a whole (and similarly for its part (c1)) could be applied pairwise to histories with the set of internal processes, say I, as parameter: $h_1 \not\in_I h_2$. However, we prefer to pair histories without such a parameter. Ideally, we would like to combine state-history pairs (states are united, histories merged) for which the histories are real-time consistent and *after that* apply the check (c1). This approach is unfeasible, as is shown by the programs

$$(P_1 :: P_2!\,0 \,\|\, P_2 :: P_1?\,x)$$

and

$$(P_1 :: (P_{11} :: P_2!\,0 \,\|\, P_{12} :: P_2!\,0) \,\|\, P_2 :: x := 0).$$

If we would follow the strategy above, the meanings of these programs would both contain the history $\langle [\langle 1, 2, 0\rangle^2] \rangle$. The problem is, that we somehow must remove this history from the meaning of the second program (it deadlocks), but reduce the same history to $\langle [\ ] \rangle$ in the meaning of the first one (showing a successful internal communication); this is clearly an impossibility.

There is, however, an easy trick to circumvent this problem. The above example suggests that we should *subtract equal communication claim records from each other* while merging: for the first program this would result in no $\langle 1, 2, 0\rangle$-records at all, while for the second program the two $\langle 1, 2, 0\rangle$-records would still be maintained. Check (c1) can then be completed by testing whether after this special merging there are any "internal communications" left, i.e., communication claims $\langle i, j, v\rangle$ with i and j internal. Formally, for $X \in \Sigma H$ and $I \in \mathbf{P}(\mathbb{N})$ we define

$$\not\in_I^{\mathrm{IC}}(X) = X \setminus \{\langle \sigma, h\rangle \mid \exists B \ll h\, \exists i, j \in I\, \exists v \in V \langle i, j, v\rangle \in B\}.$$

LEMMA. *$\not\in_I^{\mathrm{IC}}(X) \in \Sigma H$ for all $X \in \Sigma H$ and $I \in \mathbf{P}(\mathbb{N})$.*

Proof. $\not\in_I^{\mathrm{IC}}(X)$ non-empty: $X \in \Sigma H$ implies $\langle \bot, \lambda\rangle \in X$ and because there does not exist a $B \ll \lambda$ it follows that $\langle \bot, \lambda\rangle \in \not\in_I^{\mathrm{IC}}(X)$.

$\not\in_I^{\mathrm{IC}}(X)$ prefix-closed: $\not\in_I^{\mathrm{IC}}(X)$ deletes pairs from X for which the history has a certain property. Immediately from the definition it follows that all extensions of a history with this property also have this property. Reversing this we get: if a history does not have this property, then none of its prefixes can have this property. This is used in the last step of the chain of implications $\langle \sigma, h\rangle \in \not\in_I^{\mathrm{IC}}(X) \Rightarrow \langle \sigma, h\rangle \in X \Rightarrow \langle \bot, h'\rangle \in X \Rightarrow \langle \bot, h'\rangle \in \not\in_I^{\mathrm{IC}}(X)$ for all $h' \leqslant h$. ∎

The above mentioned special merge is denoted by $\#$ and does the following. Up to the length of the shortest history, $\#$ substracts equal

records (of course taking the absolute value). It is unnecessary to check especially for communication claim records because histories with equal $\langle i, j\rangle$-pairs were previously removed in the real-time consistency check. After the length of the shortest history, the longer history is just copied.

Formally: Let $h_1, h_2 \in H$. Then $h_1 \# h_2 = \langle B_k^{h_1,h_2}\rangle_{k=1}^{\max\{|h_1|,|h_2|\}}$, where $B_k^{h_1,h_2} \in \mathbf{B}(\mathrm{CAR})$ are defined as follows:

$$\text{for} \quad 1 \leqslant k \leqslant \min\{|h_1|, |h_2|\}, \qquad B_k^{h_1,h_2}(r) = |h_1(k)(r) - h_2(k)(r)|,$$

$$\text{for} \quad \min\{|h_1|, |h_2|\} < k \leqslant \max\{|h_1|, |h_2|\},$$

$$B_k^{h_1,h_2}(r) = \begin{cases} h_1(k)(r) & \text{if} \quad |h_1| > |h_2| \\ h_2(k)(r) & \text{if} \quad |h_2| > |h_1|. \end{cases}$$

In general $\#$ is commutative but not associative. However, in the context of $((T_1 \| T_2) \| T_3)$ and $(T_1 \| (T_2 \| T_3))$ we may assume because of the syntactic restriction that the visible subprocesses must be disjoint in a parallel composition (in that case vsp and VS coincide): $\mathrm{VS}(T_i) \cap \mathrm{VS}(T_j) = \varnothing$ for $1 \leqslant i < j \leqslant 3$. In that case, for $s \in S$, $\langle \sigma_i, h_i\rangle \in \mathbf{M}[\![T_i]\!]s$ $(1 \leqslant i \leqslant 3)$, it always holds that $(h_1 \# h_2) \# h_3 = h_1 \# (h_2 \# h_3)$ (see the Corollary in Section 7.5). This is used to prove the important property that $\mathbf{M}[\![((T_1 \| T_2) \| T_3)]\!]$ equals $\mathbf{M}[\![(T_1 \| (T_2 \| T_3))]\!]$, see the theorem in Section 7.5.

7.4.3.4. *An additional condition for combining state-history pairs.* When combining state-history pairs $\langle \sigma_i, h_i\rangle$, $1 \leqslant i \leqslant 2$, in the parallel composition of two processes, we should take care that the condition $\sigma_i = \perp \Rightarrow |h_i| \geqslant |h_{3-i}|$, $1 \leqslant i \leqslant 2$, is satisfied, i.e., that neither history that can be extended $(\sigma_i = \perp)$ is shorter than the other one. Here is why:

Consider the program fragment $(P_1 :: P_3!\,5 \| P_2 :: x := 0)$. For $s \in S$, $\langle \perp, \lambda\rangle \in \mathbf{M}[\![P_1 :: P_3!\,5]\!]s$ and $\langle s[0/x], \langle [\ ]\rangle\rangle \in \mathbf{M}[\![P_2 :: x := 0]\!]s$. If we combine these two state-history pairs without the extra condition above, we get the combined pair $\langle \perp, \langle [\ ]\rangle\rangle$. However, this pair should *not* belong to the parallel composition of processes P_1 and P_2, because only the internal step (the assignment) of P_2 is represented and *not* the attempt of P_1 to communicate with P_3 that occurs *at the same time.*

7.4.3.5. *The removal of real-time assumptions.* When giving the meaning of $(T_1 \| T_2)$ the real-time assumptions (represented by the no-match claim records) concerning the visible subprocesses of T_1 and T_2 should be checked. It is our policy to do this as soon as possible, that is, in the first context in which the processes i and j of a no-match claim $\langle i, j\rangle$ can be identified. The following program fragment illustrates this: $(P_1 :: P_2!\,5 \| P_2 :: P_1!\,5)$. In this case some histories of process P_1 contain the no-match claim $\langle 1, 2\rangle$ and some of process P_2, $\langle 2, 1\rangle$. After binding

processes P_1 and P_2, the real-time assumptions concerning the collection of processes $\{1, 2\}$ should be checked; in this case, exactly $\langle 1, 2\rangle$ and $\langle 2, 1\rangle$. After this check they are not needed anymore and will be removed.

In general, for $B \in \mathbf{B}(\mathrm{CAR})$ and a collection of processes $I \in \mathbf{P}(\mathbb{N})$, we can define $\mathrm{RTA}_I(B) \in \mathbf{B}(\mathrm{CAR})$ which removes from B the no-match claims concerning I:

$$\mathrm{RTA}_I(B)(r) = \begin{cases} 0 & \text{if } r = \langle i, j\rangle \text{ with } i, j \in I, \\ B(r) & \text{otherwise.} \end{cases}$$

We have to extend this operator to elements of ΣH in the same way as we extended $B[I \to i]$ to $X[I \to i]$ (see Section 7.4.2):

$$\mathrm{RTA}_I(X) = \{\langle \sigma, \langle \mathrm{RTA}_I(h(k))\rangle_{k=1}^{|h|}\rangle \mid \langle \sigma, h\rangle \in X\}.$$

LEMMA. *$\mathrm{RTA}_I(X) \in \Sigma H$ for all $X \in \Sigma H$ and $I \in \mathbf{P}(\mathbb{N})$.*

Proof. The same as for the lemma in Section 7.4.2. ∎

7.4.3.6. *Putting it all together: the meaning of* $(T_1 \parallel T_2)$.

$$\mathbf{M}[\![(T_1 \parallel T_2)]\!]s = \mathrm{RTA}_{\mathrm{tvs}}(\not\in^{\mathrm{IC}}_{\mathrm{tvs}}(\{\langle \sigma_1 \cup_s \sigma_2, h_1 \# h_2\rangle \mid \langle \sigma_i, h_i\rangle \in \mathbf{M}[\![T_i]\!]s \wedge h_1 \not\in^{\mathrm{RT}} h_2 \wedge \sigma_i = \perp \Rightarrow |h_i| \geqslant |h_{3-i}|,\ 1 \leqslant i \leqslant 2\})),$$

where $\mathrm{tvs} = \mathrm{VS}((T_1 \parallel T_2)) = \mathrm{VS}(T_1) \cup \mathrm{VS}(T_2)$, the total visible subprocesses.

LEMMA. $\{\langle \sigma_1 \cup_s \sigma_2, h_1 \# h_2\rangle \mid \langle \sigma_i, h_i\rangle \in \mathbf{M}[\![T_i]\!]\ s \wedge h_1 \not\in^{\mathrm{RT}} h_2 \wedge \sigma_i = \perp \Rightarrow |h_i| \geqslant |h_{3-i}|,\ 1 \leqslant i \leqslant 2\} \in \Sigma H$.

Proof. Abbreviate the above set to X. X non-empty: $\mathbf{M}[\![T_i]\!]s \in \Sigma H$ implies $\langle \perp, \lambda\rangle \in \mathbf{M}[\![T_i]\!]s$ $(1 \leqslant i \leqslant 2)$. $\lambda \not\in^{\mathrm{RT}} \lambda$ and $|\lambda| \geqslant |\lambda|$ are satisfied, hence $\langle \perp \cup_s \perp, \lambda \# \lambda\rangle = \langle \perp, \lambda\rangle \in X$.

X prefix-closed: Let $\langle \sigma_1 \cup_s \sigma_2, h_1 \# h_2\rangle \in X$ and $h' \leqslant h_1 \# h_2$. The proof splits into two cases, dependent on the length of h':

Case 1. $|h'| \leqslant \min\{|h_1|, |h_2|\}$. Take $h_i' \leqslant h_i$, $|h_i'| = |h'|$ $(1 \leqslant i \leqslant 2)$. Then $\langle \perp, h_i'\rangle \in \mathbf{M}[\![T_i]\!]s$ and $h_1' \not\in^{\mathrm{RT}} h_2'$ and $|h_i'| \geqslant |h_{3-i}'|$ $(1 \leqslant i \leqslant 2)$ and $h_1' \# h_2' = h'$, hence $\langle \perp \cup_s \perp, h_1' \# h_2'\rangle = \langle \perp, h'\rangle \in X$.

Case 2. $|h'| > \min\{|h_1|, |h_2|\}$. From $h' \leqslant h_1 \# h_2$ it follows that $|h'| \leqslant |h_1 \# h_2| = \max\{|h_1|, |h_2|\}$.

Taking these two conditions on $|h'|$ together we see that $|h_1| \neq |h_2|$. Without loss of generality we can suppose $|h_1| > |h_2|$. Take $h_1' \leqslant h_1$ with $|h_1'| = |h'|$. Then $\langle \perp, h_1'\rangle \in \mathbf{M}[\![T_1]\!]s$ and $\langle \sigma_2, h_2\rangle \in \mathbf{M}[\![T_2]\!]s$ and $h_1' \not\in^{\mathrm{RT}} h_2$ and $|h_1'| \geqslant |h_2|$ (and $\sigma_2 \neq \perp$ because $|h_1| > |h_2|$) and $h_1' \# h_2 = h'$, hence $\langle \perp \cup_s \sigma_2, h_1' \# h_2\rangle = \langle \perp, h'\rangle \in X$. ∎

PROPOSITION. *For all* $s \in S$, $\mathbf{M}[\![(T_1 \parallel T_2)]\!]s \in \Sigma H$.

Proof. Immediate by the lemma and the fact that both ϕ_I^{IC} and RTA_I map elements of ΣH to elements of ΣH (see the lemma in Section 7.4.3.3, respectively 7.4.3.5). ∎

7.5. *Properties of the Semantics*

In this section we derive some general properties of the semantics and use them to prove commutativity and associativity of parallel composition.

We start with a property concerning the records in the histories generated by our semantics: in the records in the histories of the semantics of a command at least one of the processes involved is a visible subprocess of that command.

HISTORY PROPERTY. *For all commands* T, $s \in S$, $\langle \sigma, h \rangle \in \mathbf{M}[\![T]\!]s$ *the following holds*:

$$\forall B \lessdot h \, \forall r \in B(\pi_1(r) \in \mathrm{VS}(T) \vee \pi_2(r) \in \mathrm{VS}(T)).$$

Proof. From the definition of $\mathbf{M}[\![T]\!]$, by an easy structural induction on T. ∎

The following lemma and its corollary concern properties in the context of the parallel composition of T_1, T_2, and T_3 (cf. the end of Section 7.4.3.3). The lemma states that under certain conditions (which are met in the case of a parallel composition) three histories cannot contain a common communication assumption record. The corollary then says that under the same conditions the special merge $\#$ of Section 7.4.3.3 is associative.

MAIN LEMMA. *Let* $s \in S$, $\langle \sigma_i, h_i \rangle \in \mathbf{M}[\![T_i]\!]s$ $(1 \leqslant i \leqslant 3)$ *and suppose* $\mathrm{VS}(T_i) \cap \mathrm{VS}(T_j) = \varnothing$ *for all* i, j, $1 \leqslant i < j \leqslant 3$. *Then for all* $r \in \mathrm{CAR}$, *all* k *such that* $1 \leqslant k \leqslant \min\{|h_i| \mid 1 \leqslant i \leqslant 3\}$ *there exists an* i, $1 \leqslant i \leqslant 3$, *with* $h_i(k)(r) = 0$.

Proof. From the History Property and the condition $\mathrm{VS}(T_i) \cap \mathrm{VS}(T_j) = \varnothing$ $(1 \leqslant i < j \leqslant 3)$ it easily follows that there cannot exist $r \in \mathrm{CAR}$ and k, $1 \leqslant k \leqslant \min\{|h_i| \mid 1 \leqslant i \leqslant 3\}$, such that $r \in h_i(k)$ for all i, $1 \leqslant i \leqslant 3$. ∎

COROLLARY. *Let* $s \in S$, $\langle \sigma_i, h_i \rangle \in \mathbf{M}[\![T_i]\!]s$ $(1 \leqslant i \leqslant 3)$ *and suppose* $\mathrm{VS}(T_i) \cap \mathrm{VS}(T_j) = \varnothing$ *for all* i, j, $1 \leqslant i < j \leqslant 3$. *Then* $(h_1 \# h_2) \# h_3 = h_1 \# (h_2 \# h_3)$.

Proof. From the Main Lemma observing that $||k-m|-n| = |k-|m-n||$ for all $k, m, n \in \mathbb{N}$ such that $k = 0$ or $m = 0$ or $n = 0$. ∎

The preceding properties enable us to prove that pairwise binding of

processes is independent of the order in which the processes are bound. E.g., for three processes $\mathbf{M}[\![((T_1 \| T_2) \| T_3)]\!]$ equals $\mathbf{M}[\![(T_1 \| (T_2 \| T_3))]\!]$. This associativity property together with commutativity $\mathbf{M}[\![(T_1 \| T_2)]\!] = \mathbf{M}[\![(T_2 \| T_1)]\!]$ justifies the writing of $\mathbf{M}[\![(T_1 \| T_2 \| T_3)]\!]$ for any order of binding T_1, T_2, and T_3. This immediately generalizes to $\mathbf{M}[\![(T_1 \| \cdots \| T_n)]\!]$ for any order of binding $T_1, ..., T_n$ $(n \geqslant 2)$.

THEOREM. $\mathbf{M}[\![(T_1 \| T_2)]\!] = \mathbf{M}[\![(T_2 \| T_1)]\!]$ *and* $\mathbf{M}[\![((T_1 \| T_2) \| T_3)]\!] = \mathbf{M}[\![(T_1 \| (T_2 \| T_3))]\!]$.

Proof. Commutativity: immediately from the commutativity of $\bigcup_s$, $\#$ and $\not\in^{\mathrm{RT}}$.

Associativity: We shall give a meaning to "$\mathbf{M}[\![(T_1 \| T_2 \| T_3)]\!]s$" and show that $\mathbf{M}[\![((T_1 \| T_2) \| T_3)]\!]s$ and $\mathbf{M}[\![(T_1 \| (T_2 \| T_3))]\!]s$ both are equal to it.

Note that in the context of the parallel composition of T_1, T_2, and T_3 (in both orders above), we may assume (see the end of Section 7.4.3.3)

(a) $\mathrm{VS}(T_i) \cap \mathrm{VS}(T_j) = \varnothing$ $(1 \leqslant i < j \leqslant 3)$.

Hence for $s \in S$, $\langle \sigma_i, h_i \rangle \in \mathbf{M}[\![T_i]\!]s$ $(1 \leqslant i \leqslant 3)$ we can apply both the Main Lemma and the Corollary. Because of associativity of $\bigcup_s$ and the Corollary we can define

$$
\begin{aligned}
&\mathbf{M}[\![(T_1 \| T_2 \| T_3)]\!]s \\
&\quad = \mathrm{RTA}_{\bigcup_{i=1}^{3} \mathrm{VS}(T_i)}(\not\in^{\mathrm{IC}}_{\bigcup_{i=1}^{3} \mathrm{VS}(T_i)} (\{\langle \sigma_1 \cup_s \sigma_2 \cup_s \sigma_3, h_1 \# h_2 \# h_3 \rangle | \\
&\qquad \langle \sigma_i, h_i \rangle \in \mathbf{M}[\![T_i]\!]s \ (1 \leqslant i \leqslant 3) \\
&\qquad \wedge\ h_i \not\in^{\mathrm{RT}} h_j \ (1 \leqslant i < j \leqslant 3) \wedge \sigma_i = \bot \Rightarrow |h_i| \geqslant |h_j| \ (1 \leqslant i, j \leqslant 3)\})).
\end{aligned}
$$

Now, for all $s \in S$,

$$
\begin{aligned}
&\mathbf{M}[\![((T_1 \| T_2) \| T_3)]\!]s \\
&\quad = \mathrm{RTA}_{\bigcup_{i=1}^{3} \mathrm{VS}(T_i)} (\not\in^{\mathrm{IC}}_{\bigcup_{i=1}^{3} \mathrm{VS}(T_i)} (\{\langle \sigma \cup_s \sigma_3, h \# h_3 \rangle | \\
&\qquad \langle \sigma, h \rangle \in \mathrm{RTA}_{\bigcup_{i=1}^{2} \mathrm{VS}(T_i)} (\not\in^{\mathrm{IC}}_{\bigcup_{i=1}^{2} \mathrm{VS}(T_i)} (\{\sigma_1 \cup_s \sigma_2, h_1 \# h_2 \rangle | \\
&\qquad \langle \sigma_1, h_1 \rangle \in \mathbf{M}[\![T_1]\!]s \wedge \langle \sigma_2, h_2 \rangle \in \mathbf{M}[\![T_2]\!]s \wedge h_1 \not\in^{\mathrm{RT}} h_2 \\
&\qquad \wedge\ \sigma_1 = \bot \Rightarrow |h_1| \geqslant |h_2| \wedge \sigma_2 = \bot \Rightarrow |h_2| \geqslant |h_1| \})) \\
&\qquad \wedge\ \langle \sigma_3, h_3 \rangle \in \mathbf{M}[\![T_3]\!]s \wedge h \not\in^{\mathrm{RT}} h_3 \wedge \sigma = \bot \Rightarrow |h| \geqslant |h_3| \\
&\qquad \wedge\ \sigma_3 = \bot \Rightarrow |h_3| \geqslant |h| \}))
\end{aligned}
$$

$$\stackrel{(*)}{=} \mathrm{RTA}_{\bigcup_{i=1}^{3} \mathrm{VS}(T_i)}(\not\in^{\mathrm{IC}}_{\bigcup_{i=1}^{3} \mathrm{VS}(T_i)}(\{\langle(\sigma_1 \cup_s \sigma_2) \cup_s \sigma_3, (h_1 \# h_2) \# h_3\rangle |$$
$$\langle\sigma_1, h_1\rangle \in \mathbf{M}[\![T_1]\!]s \wedge \langle\sigma_2, h_2\rangle \in \mathbf{M}[\![T_2]\!]s \wedge \langle\sigma_3, h_3\rangle \in \mathbf{M}[\![T_3]\!]s$$
$$\wedge\, h_1 \not\in^{\mathrm{RT}} h_2 \wedge (h_1 \# h_2) \not\in^{\mathrm{RT}} h_3 \wedge \sigma_1 = \perp \Rightarrow |h_1| \geqslant |h_2|$$
$$\wedge\, \sigma_2 = \perp \Rightarrow |h_2| \geqslant |h_1|$$
$$\wedge\, \sigma_1 \cup_s \sigma_2 = \perp \Rightarrow |h_1 \# h_2| \geqslant |h_3| \wedge \sigma_3 = \perp \Rightarrow |h_3| \geqslant |h_1 \# h_2|\}))$$
$$\stackrel{(**)}{=} \mathbf{M}[\![(T_1 \| T_2 \| T_3)]\!]s \stackrel{(***)}{=} \mathbf{M}[\![(T_1 \| (T_2 \| T_3))]\!]s,$$

where the three crucial steps are explained by

(*) We can leave out the operators $\mathrm{RTA}_{\bigcup_{i=1}^{2} \mathrm{VS}(T_i)}$ and $\not\in^{\mathrm{IC}}_{\bigcup_{i=1}^{2} \mathrm{VS}(T_i)}$ because they only concern records with $\pi_1(r) \in \bigcup_{i=1}^{2} \mathrm{VS}(T_i)$ and $\pi_2(r) \in \bigcup_{i=1}^{2} \mathrm{VS}(T_i)$. Because of the History Property and (a) it follows that such records r cannot occur in h_3. This implies that such records do not interfere with records of h_3, e.g., such records are maintained in the merge $h \# h_3$. The effect of the two above operators is then contained in the effect of $\mathrm{RTA}_{\bigcup_{i=1}^{3} \mathrm{VS}(T_i)}$ and $\not\in^{\mathrm{IC}}_{\bigcup_{i=1}^{3} \mathrm{VS}(T_i)}$, since $\bigcup_{i=1}^{2} \mathrm{VS}(T_i) \subseteq \bigcup_{i=1}^{3} \mathrm{VS}(T_i)$.

This holds because of

(1) associativity of $\bigcup_s$ and the Corollary;

(2) $(h_1 \# h_2) \not\in^{\mathrm{RT}} h_3 \Leftrightarrow (h_1 \not\in^{\mathrm{RT}} h_3 \wedge h_2 \not\in^{\mathrm{RT}} h_3)$:

$\Leftarrow$ Easy because $r \in (h_1 \# h_2)(k)$ implies that $r \in h_i(k)$ for $i = 1$ or $i = 2$.

(**) $\Rightarrow$ According to the Main Lemma, $r \in h_3(k)$ and $r \in h_i(k)$ $(1 \leqslant i \leqslant 2)$ implies that $k > |h_{3-i}|$ or that $h_{3-i}(k)(r) = 0$; in both cases $r \in (h_1 \# h_2)(k)$;

(3) $(\sigma_1 \cup_s \sigma_2) = \perp \Leftrightarrow (\sigma_1 = \perp \vee \sigma_2 = \perp)$ and $|h_1 \# h_2| = \max\{|h_1|, |h_2|\}$.

(***) The previous equations hold as well when σ_i and h_i $(1 \leqslant i \leqslant 3)$ are interchanged. ▮

7.6. *Concluding Remarks*

The proposition in Section 7.2 shows that we do not have to worry about continuity of the meaning function.

After all these technicalities the next section gives some examples which illustrate the basic ideas and what is observable.

8. EXAMPLES

In the examples below E_n abbreviates the program (fragment) of example n and s is an arbitrary element of S.

EXAMPLE 1. $E_1 \equiv P_1 :: *[\mathbf{true} \to P_2!5]$. First we compute

$$\begin{aligned}\mathbf{M}[\![[\mathbf{true} \to P_2!5]]\!]s &= \bigcup_{j=1}^{1} \mathbf{M}[\![P_2!5]\!]^*(\mathbf{G}[\![\mathbf{true}, \varnothing]\!]s) \\ &= \mathbf{M}[\![P_2!5]\!]^*(\mathrm{PFC}(\{\langle s, \lambda\rangle\})) \\ &\overset{(*)}{=} \mathbf{M}[\![P_2!5]\!]s = \mathbf{G}[\![P_2!5, \varnothing]\!]s \\ &= \mathrm{PFC}(\{\langle s, \langle[\langle 0, 2\rangle]\rangle^t \hat{}\, \langle[\langle 0, 2, 5\rangle]\rangle\rangle \mid t \in \mathbb{N}\}).\end{aligned}$$

(*) In general, by writing out the definitions of Section 7.2, we see that $\phi^*(\mathrm{PFC}(\{\langle s, \lambda\rangle\})) = \phi(s)$.

Then

$$\mathbf{M}[\![*[\mathbf{true} \to P_2!5]]\!]s = \bigcup_{i \in \mathbb{N}} \phi_i(s),$$

where

$$\phi_0(s) = \{\langle \perp, \lambda\rangle\},\ \phi_{i+1}(s) = \phi_i^*(\mathbf{M}[\![[\mathbf{true} \to P_2!5]]\!]s).$$

By induction we can prove for all $n \in \mathbb{N}$,

$$\begin{aligned}\phi_n(s) = \mathrm{PFC}(\{\langle \perp, \langle[\langle 0, 2\rangle]\rangle^{t_1} \hat{}\, \langle[\langle 0, 2, 5\rangle]\rangle \hat{}\, \cdots \hat{}\, \langle[\langle 0, 2\rangle]\rangle^{t_n} \hat{}\, \\ \langle[\langle 0, 2, 5\rangle]\rangle\rangle \mid t_1, \ldots, t_n \in \mathbb{N}\}).\end{aligned}$$

Hence

$$\begin{aligned}\mathbf{M}[\![E_1]\!]s = \mathrm{PFC}(\{\langle \perp, \langle[\langle 1, 2\rangle]\rangle^{t_1} \hat{}\, \langle[\langle 1, 2, 5\rangle]\rangle \hat{}\, \cdots \hat{}\, \langle[\langle 1, 2\rangle]\rangle^{t_n} \hat{}\, \\ \langle[\langle 1, 2, 5\rangle]\rangle\rangle \mid n \in \mathbb{N}, t_1, \ldots, t_n \in \mathbb{N}\}).\end{aligned}$$

Remark. This example shows why elements of ΣH should be non-empty. Otherwise $\varnothing$ would be the least element of ΣH and for the ϕ_i above we would then get $\phi_n(s) = \varnothing$ for all $n \in \mathbb{N}$ and hence $\mathbf{M}[\![E_1]\!]s = \varnothing$. This is caused by the fact that we should have a starting point for the histories and $\varnothing$ contains no histories at all.

EXAMPLE 2. $E_2 \equiv P_1 :: (P_{11} :: *[P_2!5 \to \mathbf{wait}\ 1 \ \square\ \mathbf{wait}\ 1 \to \mathbf{wait}\ 1] \parallel P_{12} :: \mathbf{wait}1; *[P_2!5 \to \mathbf{wait}\ 1 \ \square\ \mathbf{wait}\ 1 \to \mathbf{wait}\ 1])$.

We should have $\mathbf{M}[\![E_1]\!] = \mathbf{M}[\![E_2]\!]$! E_1 and E_2 have indeed the same observable behaviour: they both continuously try to output value 5 to

process 2. Let C abbreviate $[P_2!5 \to \textbf{wait } 1 \,\square\, \textbf{wait } 1 \to \textbf{wait } 1]$ (then $E_2 \equiv P_1 :: (P_{11} :: *C \parallel P_{12} :: \textbf{wait } 1; *C)$).

We first compute

$$\begin{aligned}
\mathbf{M}[\![C]\!]s &= \mathbf{M}[\![\textbf{wait } 1]\!]^* (\mathbf{G}[\![P_2!5, \{\textbf{wait } 1\}]\!]s) \\
&\quad \cup \mathbf{M}[\![\textbf{wait } 1]\!]^* (\mathbf{G}[\![\textbf{wait } 1, \{P_2!5\}]\!]s) \\
&= \mathbf{M}[\![\textbf{wait } 1]\!]^* (\mathrm{PFC}(\{\langle s, \langle [\langle 0, 2\rangle]\rangle^t \frown \\
&\quad \langle [\langle 0, 2, 5\rangle]\rangle\rangle \mid 0 \leqslant t < 1\})) \\
&\quad \cup \mathbf{M}[\![\textbf{wait } 1]\!]^* (\mathrm{PFC}(\{\langle s, \langle [\langle 0, 2\rangle]\rangle^1\rangle\})) \\
&= \mathrm{PFC}(\{\langle s, \langle [\langle 0, 2, 5\rangle], [\,]\rangle\rangle\}) \\
&\quad \cup \mathrm{PFC}(\{\langle s, \langle [\langle 0, 2\rangle], [\,]\rangle\rangle\}).
\end{aligned}$$

Then

$$\mathbf{M}[\![*C]\!]s = \bigcup_{i \in \mathbb{N}} \phi_i(s), \qquad \text{where } \phi_0(s) = \{\langle \perp, \lambda\rangle\},\ \phi_{i+1}(s) = \phi_i^*(\mathbf{M}[\![C]\!]s).$$

By induction we can prove for all $n \in \mathbb{N}$,

$\phi_n(s) = \mathrm{PFC}(\{\langle \perp, \langle [r_1], [\,]\rangle \frown \cdots \frown \langle [r_n], [\,]\rangle\rangle \mid$

$\forall i, 1 \leqslant i \leqslant n, r_i = \langle 0, 2, 5\rangle \vee r_i = \langle 0, 2\rangle\})$.

Hence $\mathbf{M}[\![P_{11} :: *C]\!]s = \mathrm{PFC}(\{\langle \perp, \langle [r_1], [\,]\rangle \frown \cdots \frown \langle [r_n], [\,]\rangle\rangle \mid$

$n \in \mathbb{N}, \forall i, 1 \leqslant i \leqslant n, r_i = \langle 11, 2, 5\rangle \vee r_i = \langle 11, 2\rangle\})$ and

$$\begin{aligned}
&\mathbf{M}[\![P_{12} :: \textbf{wait } 1; *C]\!]s \\
&\quad = (\mathbf{M}[\![*C]\!]^* (\mathbf{M}[\![\textbf{wait } 1]\!]s))[\{0\} \to 12] \\
&\quad = \mathrm{PFC}(\{\langle \perp, \langle [\,]\rangle \frown \langle [r_1], [\,]\rangle \frown \cdots \frown \langle [r_n], [\,]\rangle\rangle \mid \\
&\qquad n \in \mathbb{N}, \forall i, 1 \leqslant i \leqslant n, r_i = \langle 12, 2, 5\rangle \vee r_i = \langle 12, 2\rangle\}).
\end{aligned}$$

Next we compute the parallel composition of $P_{11} :: *C$ and $P_{12} :: \textbf{wait } 1; *C$:

$$\begin{aligned}
&\mathbf{M}[\![(P_{11} :: *C \parallel P_{12} :: \textbf{wait } 1; *C)]\!]s \\
&\quad = \mathrm{RTA}_{\{11,12\}}(\not\in^{\mathrm{IC}}_{\{11,12\}}(\{\langle \sigma_1 \cup_s \sigma_2, h_1 \,\#\, h_2\rangle \mid \\
&\qquad \langle \sigma_1, h_1\rangle \in \mathbf{M}[\![P_{11} :: *C]\!]s \wedge \langle \sigma_2, h_2\rangle \in \mathbf{M}[\![P_{12} :: \textbf{wait } 1; *C]\!]s \\
&\qquad \wedge h_1 \not\in^{\mathrm{RT}} h_2 \wedge \sigma_i = \perp \Rightarrow |h_i| \geqslant |h_{3-i}|, 1 \leqslant i \leqslant 2\})) \\
&\quad = \{\langle \perp, \langle [r_i]\rangle_{i=1}^n\rangle \mid n \in \mathbb{N}, \forall i, 1 \leqslant i \leqslant n, \\
&\qquad \mathrm{odd}(i) \Rightarrow (r_i = \langle 11, 2, 5\rangle \vee r_i = \langle 11, 2\rangle) \\
&\qquad \wedge \mathrm{even}(i) \Rightarrow (r_i = \langle 12, 2, 5\rangle \vee r_i = \langle 12, 2\rangle)\}.
\end{aligned}$$

Hence $\mathbf{M}[\![E_2]\!]s = \{\langle \perp, \langle [r_i] \rangle_{i=1}^{n} \rangle \mid n \in \mathbb{N},\ \forall i,\ 1 \leqslant i \leqslant n,\ r_i = \langle 1, 2, 5 \rangle \vee r_i = \langle 1, 2 \rangle\}$.

That $\mathbf{M}[\![E_1]\!] = \mathbf{M}[\![E_2]\!]$ holds, can be easily seen by an analogy with formal language theory: $\text{Prefixes}((b^*a)^*) = (a \cup b)^*$.

Remark. These two examples illustrate that established deadlock is just a special case of an infinite computation (and is not distinguishable from other infinite computations such as divergence; see the end of Section 6.2): E_1 deadlocks when process P_2 from some point on does not ask for a value to be input from process P_1; in the same context E_2 behaves more or less as "busy waiting," which is another form of infinite computation.

EXAMPLE 3. $E_3 \equiv (P_1 :: (P_{11} :: P_2!3 \parallel P_{12} :: P_2!7) \parallel P_2 :: P_1?x)$. We should get an infinite computation; in this case an established deadlock of either P_{11} or P_{12} after the successful communication of the other with P_2. First compute

$$\mathbf{M}[\![P_{11} :: P_2!3]\!]s = \text{PFC}(\{\langle s, \langle [\langle 11, 2 \rangle]\rangle^t \frown \langle [\langle 11, 2, 3 \rangle]\rangle\rangle \mid t \in \mathbb{N}\}),$$

$$\mathbf{M}[\![P_{12} :: P_2!7]\!]s = \text{PFC}(\{\langle s, \langle [\langle 12, 2 \rangle]\rangle^t \frown \langle [\langle 12, 2, 7 \rangle]\rangle\rangle \mid t \in \mathbb{N}\}),$$

and

$$\mathbf{M}[\![P_2 :: P_1?x]\!]s = \text{PFC}(\{\langle s[v/x], \langle [\langle 1, 2 \rangle]\rangle^t \frown \langle [\langle 1, 2, v \rangle]\rangle\rangle \mid v \in V,\ t \in \mathbb{N}\}).$$

Next

$$\begin{aligned}
&\mathbf{M}[\![(P_{11} :: P_2!3 \parallel P_{12} :: P_2!7)]\!]s \\
&\quad = \text{RTA}_{\{11,12\}}(\not\in^{\text{IC}}_{\{11,12\}}(\{\langle \sigma_1 \cup_s \sigma_2, h_1 \,\#\, h_2 \rangle \mid \\
&\qquad \langle \sigma_1, h_1 \rangle \in \mathbf{M}[\![P_{11} :: P_2!3]\!]s \wedge \langle \sigma_2, h_2 \rangle \in \mathbf{M}[\![P_{12} :: P_2!7]\!]s \\
&\qquad \wedge h_1 \not\in^{\text{RT}} h_2 \wedge \sigma_i = \perp \Rightarrow |h_i| \geqslant |h_{3-i}|,\ 1 \leqslant i \leqslant 2\})) \\
&\quad = \text{PFC}(\{\langle s, \langle [\langle 11, 2 \rangle, \langle 12, 2 \rangle]\rangle^{t_1} \frown \langle [\langle 11, 2, 3 \rangle, \langle 12, 2 \rangle]\rangle \frown \\
&\qquad \langle [\langle 12, 2 \rangle]\rangle^{t_2} \frown \langle [\langle 12, 2, 7 \rangle]\rangle\rangle \mid t_1, t_2 \in \mathbb{N}\} \\
&\qquad \cup \{\langle s, \langle [\langle 11, 2 \rangle, \langle 12, 2 \rangle]\rangle^t \frown \langle [\langle 11, 2, 3 \rangle, \langle 12, 2, 7 \rangle]\rangle\rangle \mid t \in \mathbb{N}\} \\
&\qquad \cup \{\langle s, \langle [\langle 11, 2 \rangle, \langle 12, 2 \rangle]\rangle^{t_1} \frown \langle [\langle 11, 2 \rangle, \langle 12, 2, 7 \rangle]\rangle \frown \\
&\qquad \langle [\langle 11, 2 \rangle]\rangle^{t_2} \frown \langle [\langle 11, 2, 3 \rangle]\rangle\rangle \mid t_1, t_2 \in \mathbb{N}\}).
\end{aligned}$$

Hence

$$\begin{aligned}
&\mathbf{M}[\![P_1 :: (P_{11} :: P_2!3 \parallel P_{12} :: P_2!7)]\!]s \\
&\quad = \mathrm{PFC}(\{\langle s, \langle[\langle 1,2\rangle^2]\rangle^{t_1} \frown \langle[\langle 1,2,3\rangle, \langle 1,2\rangle]\rangle \frown \\
&\qquad \langle[\langle 1,2\rangle]\rangle^{t_2} \frown \langle[\langle 1,2,7\rangle]\rangle\rangle \mid t_1, t_2 \in \mathbb{N}\} \\
&\qquad \cup \{\langle s, \langle[\langle 1,2\rangle^2]\rangle^{t} \frown \langle[\langle 1,2,3\rangle, \langle 1,2,7\rangle]\rangle\rangle \mid t \in \mathbb{N}\} \\
&\qquad \cup \{\langle s, \langle[\langle 1,2\rangle^2]\rangle^{t_1} \frown \langle[\langle 1,2\rangle, \langle 1,2,7\rangle]\rangle \frown \\
&\qquad \langle[\langle 1,2\rangle]\rangle^{t_2} \frown \langle[\langle 1,2,3\rangle]\rangle\rangle \mid t_1, t_2 \in \mathbb{N}\}).
\end{aligned}$$

Note that here the use of bags instead of sets is essential, especially if we replace 3 and 7 both by the same value. Then

$$\begin{aligned}
\mathbf{M}[\![E_3]\!]s &= \mathrm{RTA}_{\{1,2\}}(\mathfrak{C}^{\mathrm{IC}}_{\{1,2\}}(\{\langle \sigma_1 \cup_s \sigma_2, h_1 \# h_2\rangle \mid \\
&\qquad \langle \sigma_1, h_1\rangle \in \mathbf{M}[\![P_1 :: (P_{11} :: P_2!3 \parallel P_{12} :: P_2!7)]\!]s \\
&\qquad \wedge \langle \sigma_2, h_2\rangle \in \mathbf{M}[\![P_2 :: P_1?x]\!]s \wedge h_1 \not\mathfrak{C}^{\mathrm{RT}} h_2 \\
&\qquad \wedge \sigma_i = \perp \Rightarrow |h_i| \geqslant |h_{3-i}|, 1 \leqslant i \leqslant 2\})) \\
&= \mathrm{RTA}_{\{1,2\}}(\mathrm{PFC}(\{\langle \perp, \langle[\langle 1,2\rangle]\rangle \frown \langle[\langle 1,2\rangle]\rangle^{t_2}\rangle \mid t_2 \in \mathbb{N}\})) \\
&= \{\langle \perp, \langle[\,]\rangle^{t}\rangle \mid t \in \mathbb{N}\}.
\end{aligned}$$

EXAMPLE 4. $E_4 \equiv (P_1 :: (P_{11} :: P_2!3 \parallel P_{12} :: P_2!7) \parallel P_2 :: P_1?x; P_1?x)$. In this example one of the processes P_{11} and P_{12} first communicates with P_2 and then the other. The total program terminates in two time units. For $\mathbf{M}[\![P_1 :: (P_{11} :: P_2!3 \parallel P_{12} :: P_2!7)]\!]s$, see Example 3. Furthermore,

$$\begin{aligned}
&\mathbf{M}[\![P_2 :: P_1?x; P_1?x]\!]s \\
&\quad = \mathrm{PFC}(\{\langle s[v_2/x], \langle[\langle 1,2\rangle]\rangle^{t_1} \frown \langle[\langle 1,2,v_1\rangle]\rangle \frown \\
&\qquad \langle[\langle 1,2\rangle]\rangle^{t_2} \frown \langle[\langle 1,2,v_2\rangle]\rangle\rangle \mid v_1, v_2 \in V, t_1, t_2 \in \mathbb{N}\}).
\end{aligned}$$

Then

$$\begin{aligned}
\mathbf{M}[\![E_4]\!]s &= \mathrm{RTA}_{\{1,2\}}(\mathrm{PFC}(\{\langle s[7/x], \langle[\langle 1,2\rangle], [\,]\rangle\rangle\} \\
&\qquad \cup \{\langle s[3/x], \langle[\langle 1,2\rangle], [\,]\rangle\rangle\})) \\
&= \mathrm{PFC}(\{\langle s[v/x], \langle[\,]\rangle^2\rangle \mid v \in \{3,7\}\}).
\end{aligned}$$

EXAMPLE 5. $E_5 \equiv (P_1 :: (P_{11} :: P_2!3 \parallel P_{12} :: P_2!7) \parallel P_2 :: (P_{21} :: P_1?x \parallel P_{22} :: P_1?y))$. In this example processes P_{11} and P_{12} communicate *simultaneously* with processes P_{21} and P_{22}. The total program terminates in

one time unit. For $\mathbf{M}[\![P_1 :: (P_{11} :: P_2!3 \parallel P_{12} :: P_2!7)]\!]s$, see Example 3. Similarly we can compute

$$\begin{aligned}
&\mathbf{M}[\![P_2 :: (P_{21} :: P_1?x \parallel P_{22} :: P_1?y)]\!]s \\
&\quad = \mathrm{PFC}(\{\langle s[v_1/x][v_2/y], \langle[\langle 1,2\rangle^2]\rangle^{t_1} \hat{}\ \langle[\langle 1,2,v_1\rangle, \langle 1,2\rangle]\rangle \hat{}\ \\
&\qquad \langle[\langle 1,2\rangle]\rangle^{t_2} \hat{}\ \langle[\langle 1,2,v_2\rangle]\rangle\rangle \mid v_1, v_2 \in V, t_1, t_2 \in \mathbb{N}\} \\
&\qquad \cup \{\langle s[v_1/x][v_2/y], \langle[\langle 1,2\rangle^2]\rangle^{t} \hat{}\ \\
&\qquad \langle[\langle 1,2,v_1\rangle, \langle 1,2,v_2\rangle]\rangle\rangle \mid v_1, v_2 \in V, t \in \mathbb{N}\} \\
&\qquad \cup \{\langle s[v_1/x][v_2/y], \langle[\langle 1,2\rangle^2]\rangle^{t_1} \hat{}\ \langle[\langle 1,2\rangle, \langle 1,2,v_2\rangle]\rangle \hat{}\ \\
&\qquad \langle[\langle 1,2\rangle]\rangle^{t_2} \hat{}\ \langle[\langle 1,2,v_1\rangle]\rangle\rangle \mid v_1, v_2 \in V, t_1, t_2 \in \mathbb{N}\}).
\end{aligned}$$

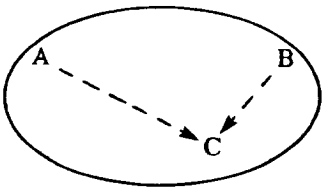
Then

$$\begin{aligned}
\mathbf{M}[\![E_5]\!]s &= \mathrm{RTA}_{\{1,2\}}(\mathrm{PFC}(\{\langle s[v_1/x][v_2/y], \langle[\,]\rangle\rangle \mid (v_1 = 3 \wedge v_2 = 7) \\
&\qquad \vee (v_1 = 7 \wedge v_2 = 3)\})) \\
&= \mathrm{PFC}(\{\langle s[v_1/x][v_2/y], \langle[\,]\rangle\rangle \mid (v_1 = 3 \wedge v_2 = 7) \\
&\qquad \vee (v_1 = 7 \wedge v_2 = 3)\}).
\end{aligned}$$

9. Real-Time Models

9.1. *Introduction*

The maximal parallelism model as used here, is flawed by some conceptual problems. We illustrate these problems with an example. Consider a network with distributed control, and two processes A and B in different nodes that want to communicate with a process C in a third node. If A wants to communicate at an earlier time than B, relative to some global time scale, then according to the fcfs-principle, indeed, A should communicate first. Whether A's message *arrives* in C before B's message or not,



depends on the topology of the network. So, imposing a fcfs-principle upon the order of communications induces non-trivial requirements upon an

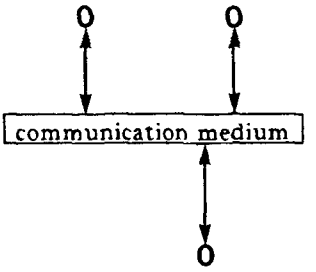
underlying communication layer; requirements that we would not like to make. Similar problems occur if processors communicate, e.g., via a common bus where assumptions about bus-arbitration have to be taken into account.

The lesson that should be drawn from this example is, that whereas our current model applies the fcfs-principle to the order of initiations of requests, the principle should rather be applied to the order in which a process becomes aware of requests. In doing so, we create the freedom to relax the stringent impositions of the original model on the behaviour of a communication layer. Specifically, in this way it becomes possible to vary the time gap (0 in the original model) between the initiation and receipt of a communication request, which reflects the uncertainties about the communication layer.

This variation of the time gap is the essential feature of the $MAX_\gamma(\delta, \varepsilon)$ model of distributed concurrency. The parameters δ and ε function as lower and upper bound on the above time gaps which are allowed to take on any value inbetween these bounds. As a consequence, communications that are initiated too close in time (relative to a global clock) cannot be temporally ordered anymore. These time bounds may be interpreted as an abstraction of the propagation delays within some communication layer. The third parameter, γ, of the model is used to extend communications in time and denotes the number of time units it takes.

9.2. $MAX_\gamma(\delta, \varepsilon)$ *Model of Concurrency*

The model is based on the Salwicki–Müldner maximal parallelism model: there is no unnecessary waiting between the execution of actions. Com-



munication between processes is served on a first-come first-served basis. Additionally, the following model pertains to process-communication:

— processes communicate via a medium;

— it takes between δ and ε time units (ε *not* included) for the medium to become aware of a process expressing its willingness to communicate or withdrawing its willingness (time-out);

— communication between two processes only occurs after the medium has become aware of both processes' willingness;

— a communication takes an additional γ time units during which period the processes remain synchronized;

— a communication that is in progress at a time when the medium receives a time-out from one of the participating processes, will be completed; a communication that might be started at such a time, will not be executed.

Remarks. — Communication always takes at least $\delta + \gamma$ time units.

— $\mathrm{MAX}_0(0, 2) \not\models \{\mathrm{true}\}(P_1 :: P_2?x;\ P_2?y \ \|\ P_2 :: (P_{21} :: P_1!1 \ \|\ P_{22} :: \textbf{wait}\ 1;\ P_1!2))\{x = 1\}$, and $\mathrm{MAX}_0(0, 1) \models \{\mathrm{true}\}(P_1 :: P_2?x;\ P_2?y \ \|\ P_2 :: (P_{21} :: P_1!1 \ \|\ P_{22} :: \textbf{wait}\ 1;\ P_1!2))\{x = 1\}$. In other words, there is an uncertainty interval of $\varepsilon - \delta$: if requests for communications are initiated $\varepsilon - \delta$ or more time units apart, the first request will indeed be served first; if, on the other hand, these requests are initiated within this interval, the order in which these requests are served is undefined.

— $\mathrm{MAX}_0(0, 1)$ gives rise to pure maximal parallelism; $\mathrm{MAX}_0(0, \infty)$ to pure interleaving semantics (with respect to the communication actions). It is to have the latter correspondence that the medium has to become aware of requests *within* ε time units. Otherwise, $\mathrm{MAX}_0(0, \infty)$ would allow infinite delays.

10. Real-Time Semantics for Mini CSP-R

The $\mathrm{MAX}_\gamma(\delta, \varepsilon)$ model only influences the semantics of communication actions. So, the definition of the auxiliary function **G** has to change, but no additional changes are needed in the definition of **M**. The intention of these changes is to have **G** "generate" any history that is consistent with the parameters of the model. As these (additional) consistency requirements are a purely local affair, the parallel composition of processes requires no additional effort.

Consider an I/O command. The changes in the sets of generated histories that $\mathrm{MAX}_\gamma(\delta, \varepsilon)$ induces are three-fold:

1. Histories must be generated in which the first waiting-action (i.e., the first no-match claim record) ocurs τ time units later than the time at which communication was requested; and this for any τ such that $\delta \leqslant \tau < \varepsilon$.

2. In no history can communication or waiting start within δ time units of the request.

3. Communication takes γ time units; this is modelled by having the

associated communication claim record mark the time, in a history, at which communication starts and by appending empty bags to trace out γ time units.

The changes to **G** are complicated by the necessity of applying the above considerations to every I/O command in the environment (i.e., in the selection or repetition).

Hence, to take care of the first point above, the basic idea is to associate with an environment $\{g_1, ..., g_n\}$ a set of times $\{t_1, ..., t_n\}$ such that $\delta \leqslant t_i < \varepsilon$. These times represent the delays of the first waiting action for the corresponding guards, i.e., the delays until the medium becomes aware of the corresponding requests. One such choice corresponds to one possible history. To generate the corresponding sequences of bags of no-match claim records, we introduce two auxiliary functions:

DEFINITION. For sets of guards G and times (i.e., natural numbers) T, time t, and state $s \in S$, define

— $A(G, T, t) = \{g_i \in G \mid t_i < t, 1 \leqslant i \leqslant n\}$, where $\{g_1, ..., g_n\} \subseteq G$ are the I/O guards in G and $T = \{t_1, ..., t_n\}$,

— $\mathrm{Ext}(G, T, t, s) = \langle \mathrm{RTA}(A(G, T, k), s) \rangle_{k=1}^{t}$.

$\mathrm{Ext}(\{\alpha_1, ..., \alpha_n\}, \{t_1, ..., t_n\}, t, s)$ yields a sequence of bags of no-match claim records for the I/O commands $\alpha_1, ..., \alpha_n$. The time t_i represents the delay of the first waiting action (i.e., no-match claim record) for α_i; t is the time at which communication or a time-out occurs. The function A is auxiliary to Ext.

Now, we are ready to define **G** (terminology as in Section 7):

$$\mathbf{G}[\![b, A]\!]s = \begin{cases} \mathrm{PFC}(\{\langle s, \lambda \rangle\}) & \text{if } \mathbf{W}[\![b]\!]s, \\ \{\langle \perp, \lambda \rangle\} & \text{otherwise.} \end{cases}$$

$$\mathbf{G}[\![\textbf{wait } d, A]\!]s = \mathrm{PFC}(\{\langle s, \mathrm{Ext}(A, T, t+\tau, s) \rangle \mid \max\{\mathbf{V}[\![d]\!]s, 1\}$$

$$= \mathrm{minwait}(A \cup \{\textbf{wait } d\}, s) \stackrel{\mathrm{def}}{=} t,$$

$$T \stackrel{\mathrm{def}}{=} \{t_1, ..., t_n\}, \delta \leqslant t_i < \varepsilon (1 \leqslant i \leqslant n), \delta \leqslant \tau < \varepsilon\}).$$

where n is the number of I/O guards in A.

$$\mathbf{G}[\![P_j!e, A]\!]s = \mathrm{PFC}(\{\langle s, \mathrm{Ext}(\mathrm{GRDS}, T, t, s) \hat{}\, \langle [\langle 0, j, \mathbf{V}[\![e]\!]s \rangle] \rangle \hat{}\, \langle [\,] \rangle^{\gamma} \rangle$$

$$\delta \leqslant t < \mathrm{minwait}(A, s) + \varepsilon - 1,$$

$$T \stackrel{\mathrm{def}}{=} \{t_1, ..., t_n\}, \delta \leqslant t_i < \varepsilon\ (1 \leqslant i \leqslant n)\}),$$

where $\mathrm{GRDS} = A \cup \{P_j!e\}$ and n is the number of I/O guards in GRDS.

The upperbound on t takes the delay of the arrival of the time-out message in the medium into account. The "-1" factor corresponds to the fact that the medium becomes aware of requests *before* ε time units have elapsed.

$$\mathbf{G}[\![P_j?x, A]\!]s = \mathrm{PFC}(\{\langle s[v/x], \mathrm{Ext}(\mathrm{GRDS}, T, t, s) \mathbin{\hat{}} \langle[\langle j, 0, v\rangle]\rangle \mathbin{\hat{}} \langle[\,]\rangle^{\gamma}\rangle$$
$$v \in V,\ \delta \leqslant t < \mathrm{minwait}(A, s) + \varepsilon - 1,$$
$$T \stackrel{\mathrm{def}}{=} \{t_1, \ldots, t_n\},\ \delta \leqslant t_i < \varepsilon (1 \leqslant i \leqslant n)\}),$$

where $\mathrm{GRDS} = A \cup \{P_j?x\}$ and n is the number of I/O guards in GRDS.

$$\mathbf{G}[\![b; g, A]\!]s = \mathbf{G}[\![g, A]\!]^{*}(\mathbf{G}[\![b, A]\!]s),$$

where g is either a pure I/O guard or a pure wait guard.

We illustrate these equations by the example in the second remark of Section 9.2. Let

$$P \equiv (P_1 :: P_2?x;\ P_2?y \parallel P_2 :: (P_{21} :: P_1!1 \parallel P_{22} :: \textbf{wait}\ 1;\ P_1!2)).$$

We claim that $\mathrm{MAX}_0(0, 2) \not\models \{\mathrm{true}\}\ P\ \{x = 1\}$ but $\mathrm{MAX}_0(0, 1) \models \{\mathrm{true}\}\ P\ \{x = 1\}$. In other words, we claim that $\mathrm{MAX}_0(0, 2)$ allows computations in which P_{22} communicates first, that are disallowed by $\mathrm{MAX}_0(0, 1)$. So, assume $\gamma = 0$, $\delta = 0$, $\varepsilon = 2$.

$$\mathbf{M}[\![P_1!2]\!]s = \mathbf{G}[\![P_1!2, \varnothing]\!]s$$
$$= \mathrm{PFC}(\{\langle s, \mathrm{Ext}(\{P_1!2\}, \{t_1\}, t, s) \mathbin{\hat{}} \langle[\langle 0, 1, 2\rangle]\rangle \mid$$
$$t \in \mathbb{N},\ 0 \leqslant t_1 \leqslant 1\}).$$

Now, $\mathrm{Ext}(\{P_1!2\}, \{0\}, t, s) = \langle \mathrm{RTA}(\{P_1!2\}, s)\rangle^{t} = \langle[\langle 0, 1\rangle]\rangle^{t}$, $t \in \mathbb{N}$, $\mathrm{Ext}(\{P_1!2\}, \{1\}, 0, s) = \lambda$, and

$$\mathrm{Ext}(\{P_1!2\}, \{1\}, t, s) = \langle \mathrm{RTA}(\varnothing, s)\rangle \mathbin{\hat{}} \langle \mathrm{RTA}(\{P_1!2\}, s)\rangle^{t-1}$$
$$= \langle[\,]\rangle \mathbin{\hat{}} \langle[\langle 0, 1\rangle]\rangle^{t-1},\ t > 0.$$

Hence

$$\mathbf{M}[\![P_1!2]\!]s = \mathrm{PFC}(\{\langle s, \langle[\,]\rangle^{\tau} \mathbin{\hat{}} \langle[\langle 0, 1\rangle]\rangle^{t} \mathbin{\hat{}}$$
$$\langle[\langle 0, 1, 2\rangle]\rangle \mid 0 \leqslant \tau \leqslant 1,\ t \in \mathbb{N}\}).$$

Analogously, we obtain the semantics of $P_1!1$ and of the input commands of P_1. Moreover, $\mathbf{M}[\![\textbf{wait}\ 1]\!]s = \mathrm{PFC}(\{\langle s, \langle[\,]\rangle^{\tau}\rangle \mid 1 \leqslant \tau \leqslant 2\})$, hence

$$\mathbf{M}[\![P_{21} :: P_1!1]\!]s = \mathrm{PFC}(\{\langle s, \langle[\,]\rangle^{\tau_{21}} \mathbin{\hat{}} \langle[\langle 21, 1\rangle]\rangle^{t_{21}} \mathbin{\hat{}}$$
$$\langle[\langle 21, 1, 1\rangle]\rangle \mid 0 \leqslant \tau_{21} \leqslant 1,\ t_{21} \in \mathbb{N}\})$$

$$\mathbf{M}[\![P_{22} :: \textbf{wait } 1; P_1!2]\!]s = \mathrm{PFC}(\{\langle s, \langle [\,]\rangle^{\tau_{22}} \hat{}\, \langle [\langle 22, 1\rangle]\rangle^{t_{22}} \hat{}\, \langle [\langle 22, 1, 2\rangle]\rangle\rangle \mid 1 \leqslant \tau_{22} \leqslant 3,\ t_{22} \in \mathbb{N}\})$$

$$\mathbf{M}[\![P_1 :: P_2?x; P_2?y]\!]s = \mathrm{PFC}(\{\langle s[v_1/x][v_2/y], \langle [\,]\rangle^{\tau_{11}} \hat{}\, \langle [\langle 2, 1\rangle]\rangle^{t_{11}} \hat{}\, \langle [\langle 2, 1, v_1\rangle]\rangle \hat{}\, \langle [\,]\rangle^{\tau_{12}} \hat{}\, \langle [\langle 2, 1\rangle]\rangle^{t_{12}} \hat{}\, \langle [\langle 2, 1, v_2\rangle]\rangle\rangle \mid 0 \leqslant \tau_{11}, \tau_{12} \leqslant 1,\ t_{11}, t_{12} \in \mathbb{N},\ v_1, v_2 \in V\}).$$

Consider the histories for P_{21} and P_{22} in which $\tau_{21} = \tau_{22} = t_{21} = 1$, $t_{22} = 0$. In particular consider P_{21}'s history $\langle [\,], [\langle 21, 1\rangle], [\langle 21, 1, 1\rangle]\rangle$ and P_{22}'s history $\langle [\,], [\langle 22, 1, 2\rangle]\rangle$. These compatible histories yield the following history for P_2: $\langle [\,], [\langle 2, 1\rangle, \langle 2, 1, 2\rangle], [\langle 2, 1, 1\rangle]\rangle$. This is compatible with P_1's history $\langle [\,], [\langle 2, 1, 2\rangle], [\langle 2, 1, 1\rangle]\rangle$, obtained by taking $\tau_{11} = 1$, $\tau_{12} = t_{11} = t_{12} = 0$, $v_1 = 2$, $v_2 = 1$. From these two histories we can compute the following element in the denotation for P: $\langle s[2/x][1/y], \langle [\,]\rangle^3\rangle$. To show that this computation cannot be generated by the $\mathrm{MAX}_0(0, 1)$-model (i.e., the maximal parallelism model, as used in Section 7) is straightforward: now, choosing $\tau_{11} = \tau_{21} = 1$ is illegal (cf. Example 4 in Section 8).

11. Conclusions

We have given a denotational semantics for real-time distributed computing stressing:

(1) compositionality, thus supplying a basis for compositional specification and verification techniques,

(2) a model of concurrency that is realistic, in contrast with interleaving, in the context of real-time: the maximal parallelism model,

(3) simplicity by basing our techniques upon the linear history semantics for CSP of Francez *et al.* (1984).

We feel that our way of dealing with real-time is particularly simple. Timing aspects of programs relate to the length of the histories. Maximal parallelism constraints are made explicit by recording not only the occurrence of communications but also the act of waiting for one. When binding two processes, these constraints imply that at no instant of time both processes are waiting for a mutual communication.

Exact clocking of instructions is unrealistic because then all actions can be exactly determined in time. In a shared variables context, this would

imply that mutual exclusion, for example, could be programmed without any additional means such as semaphores. This is resolved in Milner's SCCS by introducing the nondeterministic but bounded wait synchronization primitive δ *which may violate* the maximal parallelism constraints. In our setup, however, shared variables are excluded, so the mutual exclusion anomaly above does not occur. Additionally, by extending the maximal parallelism model by introducing non-deterministic intervals modelling synchronization delays, again this anomaly disappears. Halpern *et al.* (1985) arrived independently at the same extended model, in their case to achieve coordinated actions in a real-time distributed system. This extension furthermore shows that our techniques can easily accomodate more detailed real-time features. Another example of this is modelling the drifting of local clocks. Since only initial and final states and histories are observable, we hope that exact clocking of instructions together with the extension of the maximal parallelism model result in a realistic simplification of the phenomena inherent in the description of real-time *distributed* computing.

We based our research on CSP-R, a language that captures the essential real-time features of Ada, as supported by the simulation of Ada by CSP-R in Appendix A2. In fact, we had to solve three problems: First, how to model maximal concurrency in a compositional way. Second, how to deal with CSPs particular form of naming communication partners, i.e., of process-naming. The latter is a non-trivial problem and its solution definitely complicates our semantics: the use of bags instead of sets in our histories and many of the complications in parallel composition are a direct consequence of it. Third, the rather peculiar semantics of Ada's delay guards, as occurring in, e.g., selective waits with delay statement **delay** 0. Our ideas about modelling maximal parallelism are independent of this and, we claim, are of general applicability. This is illustrated by (Gerth, 1985; Huizing *et al.*, 1987) in which a formal semantics for (recursive) Occam is given, that is surprisingly simple because of the much cleaner communication mechanism of Occam, using communication channels between pairs of processes.

There is a clear correspondence between the readiness semantics of CSP (see Hehner and Hoare, 1983) and ours: our sets of no-match claim records—like the ready sets—record the disposition to participate in certain communications. There is also a clear difference, since unlike ready sets, a no-match claim record witnesses such a disposition at only one time instant and does not imply anything about future behaviour. Since dispositions change over time this means that we have to record such dispositions at every time instant. There is also a difference in use since, apart from detecting deadlock, no-match claim records are also used to enforce maximal parallelism.

Certain aspects which cause the readiness model to be not fully abstract, thus leading to the failure set model (see Brookes *et al.*, 1984), are also present in our model: Our semantics differentiates the two program fragments

$$[\mathbf{true} \to P_1!0; \mathbf{wait}\,1 \;\square\; \mathbf{true} \to P_2!0; \mathbf{wait}\;1]$$

and

$$[\mathbf{true} \to P_1!0; \mathbf{wait}\;1 \;\square\; \mathbf{true} \to P_2!0; \mathbf{wait}\;1 \;\square\; \mathbf{true} \to [P_1!0 \to \mathbf{wait}\;1 \;\square\; P_2!0 \to \mathbf{wait}\;1]],$$

although their observable behaviour is the same.

In (Huizing *et al.*, 1987) the authors develop a fully abstract version of our semantics for an Occam-like language and give a proof of full abstractness. Like for the ready set semantics, full abstraction is attained by an "upward closure" operation on the no-match claim records. In Gerth and Boucher (1987) the resulting model is considerably simplified and developed as an extension of the failure set model. In fact, independently from us, Andy Boucher (1986) developed quite similar techniques to give denotational semantics to Occam.

Having discovered on a semantic level how to reason compositionally about maximal parallelism we now have a firm basis for developing compositional specification and verification methods. In fact, the present paper laid the foundation for our participation in ESPRIT project 937: Debugging and Specification of Ada Real-Time Embedded Systems (DESCARTES). In the context of this project we have applied similar techniques to obtain a fully abstract model for statecharts (Harel, 1987), a language akin to ESTEREL (Berry and Cosserat, 1985). Furthermore, we used our compositional semantics to get a compositional proof theory for a subset of CSP-R (Hooman, 1987, 1988), generalizing the work of Zwiers *et al.* (1985, 1988) to real-time.

APPENDIX A: CSP-R AND THE SIMULATION OF ADA

A1. *CSP-R*

The only difference between Mini CSP-R (see Section 2) and CSP-R lies in the definition of I/O commands. CSP-R extends Mini CSP-R in the following ways:

— Communication takes place via (a form of) channels,

— The expressions in output commands and the variables in input commands are vectors,

— Process identifiers can be communicated and can be used in subsequent communications to determine the target process,

— Communication with an arbitrary process can be requested instead of only addressing a particular process.

The syntax of Mini CSP-R is changed in the following way: Replace forms 3.1 and 3.2 of the instructions by

3.1.1. $P_i.c!\mathbf{e}$ —output to process i via channel c the values of the expressions in the list **e**, together with the identification of the sending process

3.1.2. id.$c!\mathbf{e}$ —as 3.1.1, but now the target process is determined by the value of the identification variable id.

3.1.3. $.c!\mathbf{e}[\#\mathrm{id}]$ —output via channel c to *any* process the values of the expressions in the list **e**, together with the identification of the sender; record the identity of the receiving process in the identification variable id (the brackets [and] indicate that the identification variable is optional, i.e., $.c!\mathbf{e}$ is allowed, too)

3.2.1. $P_i.c?\mathbf{x}$ —the analogon of 3.1.1, but now values are received and are assigned to the variables in the list **x**

3.2.2. id.$c?\mathbf{x}$ —the analogon of 3.1.2

3.2.3. $.c?\mathbf{x}[\#\mathrm{id}]$—the analogon of 3.1.3.

An identification variable is a variable ranging over $\{P_1, P_2, \ldots\}$. It can only be assigned to using an instruction of the form 3.1.3 or 3.2.3.

The notions of syntactic and semantic matching of I/O commands have to be reformulated. $\langle P_i, \alpha \rangle$ and $\langle P_j, \beta \rangle$ match syntactically iff:

1. α and β specify the same channel,
2. the vectors have equal length,
3. if α is an input command, then β is an output command and vice versa, and
4. if $\alpha(\beta)$ is of the form 3.1.1 or 3.2.1 then the specified target process should be $P_j(P_i)$.

$\langle i, \alpha \rangle$ and $\langle j, \beta \rangle$ match semantically iff:

1. $\langle P_i, \alpha \rangle$ and $\langle P_j, \beta \rangle$ match syntactically,
2. control in P_i and P_j is in front of both α and β, and
3. if $\alpha(\beta)$ is of the form 3.1.2 or 3.2.2, then the identification variable must have the value $P_j(P_i)$.

The result of two semantically matching I/O commands is the simultaneous execution of those commands as indicated by 3.1.1–3.2.3 above. Its effect is the assignment of the expression values to the variables and, possibly, the assignment to identification variables. Because of form 3.1.3 and 3.2.3 it is possible that $\langle i, \alpha \rangle$ has more than one semantic match $\langle j, \beta \rangle$. In that case, one of these β's is non-deterministically chosen and executed simultaneously with α.

The remaining syntax and interpretation of CSP-R is the same as for Mini CSP-R.

As for the extension of our denotational semantics to CSP-R, like the assumptions we have to record about values in the denotations for input commands, we now additionally record assumptions about the communication target in the denotations for I/O commands of the forms 3.1.3 and 3.2.3.

Of course, the communication assumption records have to change. The communication claim records now have to record the communication channel and the communicated *vector* of values (instead of a single value). The no-match claim records now record the communication channel and the *length* of the communicated vector of values. Additionally, because of the I/O commands of the forms 3.1.3 and 3.2.3, no-match claim records have to indicate with which *set* of processes a match is impossible (a single process for the forms 3.1.1, 3.1.2, 3.2.1, and 3.2.2, and all processes for the forms 3.1.3 and 3.2.3).

The denotations and techniques such as the consistency check have to be adapted corresponding to the above changes. These adaptations are straightforward except for a slight complication in the meaning of $P_i :: T$: Because any communication target is assumed in the denotations for I/O commands of the forms 3.1.3 and 3.2.3, now constructs like $P_i :: .c!\mathbf{e}$ generate communication claim records in which process i communicates with itself. This is clearly impossible and such records should be removed by an additional operator. (Notice that this problem did not occur for Mini CSP-R, because constructs like $P_i :: P_i!e$ were prohibited syntactically by the naming conventions, see Section 2.) The resulting semantics can be found in (Koymans, 1984).

A2. *Simulating Ada*

To illustrate the power of CSP-R we translate the basic Ada communication primitives into CSP-R. This translation is denoted by τ. The Ada rendezvous is assumed to be understood.

1. The timed entry call (Ada, 1983, Section 9.7.3).

   select $T_i.a(\mathbf{e}, \mathbf{x})$; S_1 **or delay** t; S_2 **end select**;

The semantics of this statement prescribes that if a rendezvous can be started within the specified duration t (or immediately), then it is performed and S_1 is executed afterwards. Otherwise, when the duration has expired, S_2 is executed.

We offer as translation

$$[T_i.a!(\mathbf{e}, \mathbf{x}) \rightarrow T_i.a?\mathbf{x}; \tau(S_1) \,\square\, \textbf{wait } t \rightarrow \tau(S_2)].$$

2. The selective wait (without terminate alternative)(Ada, 1983, Section 9.7.1).

select or$(i = 1..n)$ **when** $b_i \Rightarrow S_i$ **or** $(j = 1..m)$
when $b^j \Rightarrow$ **delay** E^j; S^j **end select**;
where $S_i \equiv$ **accept** $a_i(\mathbf{u}_i \# \mathbf{v}_i)$ **do** S_{i_1} **end**; S_{i_2} $(i = 1..n)$.

The semantics is, that first the minimum value MIN, of those E^j whose guard, b^j, is open is evaluated. If a rendezvous with one of the a_i's whose guard, b_i, is open, can be started either immediately or within duration MIN, then it is performed and S_{i_2} is executed afterwards. Otherwise, when MIN time units have elapsed, one of the delay alternatives S^j for which $E^j = \text{MIN}$ (and whose associated guard is open) is executed.

Our translation:

$$\left[\,\square_{i=1}^{n}\, b_i; .a_i?(\mathbf{u}_i, \mathbf{v}_i) \# \text{id} \rightarrow \tau(S_{i_1}); \text{id}.\, a_i!\mathbf{v}_i; \tau(S_{i_2}) \,\square\, \square_{j=1}^{m}\, b^j; \textbf{wait } E^j \rightarrow \tau(S^j)\right].$$

We quote (Ada, 1983, Section 9.7.1) for the semantics of a delay alternative in a selective wait: "an open delay alternative will be selected if no accept alternative can be selected before the specified delay has elapsed (immediately, for a negative or zero delay in the absence of queued entry calls)." This means that a delay alternative **delay** 0 is selected *immediately*, although it should be checked whether there are no queued entry calls. Not only is this unrealistic, it also gives rise to the following anomaly: Consider a call of the recursive procedure P declared by

procedure $P =$ **begin select accept** A; **or delay** 0; P; **end select end**;

in a context where entry A is not called immediately. According to (Ada, 1983) there need not pass any time between the calling of P and any inner call of P, i.e., an infinite execution sequence takes no execution time!

Note that we could incorporate recursion easily into CSP-R on account of the structure of our semantic domain. Anyway, even in CSP-R without recursion, we can expand the calling of P arbitrarily deep. Keeping the same semantics as in (Ada, 1983) would then mean that an arbitrarily long execution sequence would take no execution time.

We removed this anomaly in our semantics by making **wait** 0 equivalent to **wait** 1 (that is, a wait guard has a waitvalue of at least 1, see Sections 2 and 7), thus reflecting the fact that it takes time to check whether immediate communication is possible or not. Now we get the desired semantics by simply translating Ada's **delay** t into CSP-R's **wait** t.

It is interesting to note that our techniques are in fact not capable of modelling the anomaly above: In our semantics the assumptions on the impossibility of communication are incorporated *within* the history, in fact within the mechanism that describes the passage of time. If we would have formulated these assumptions as independent conditions *on* the history (which would then contain only communication claim records), the modelling of the above anomaly would have been possible. E.g., when calling procedure P above, an *empty communication history* is produced *under the condition* that entry A is not called immediately. Such independent conditions, however, would disturb the simple structure of our semantic domain and for such an unrealistic possibility in the Ada semantics this is certainly not worth the trouble.

Appendix B: Definition of $\mathbf{B}[I \to i]$

Definition 1. For $I, J \in \mathbf{P}(\mathbb{N})$ define $R(I, J) \in \mathbf{P}(\mathrm{CAR})$ as $R(I, J) = \{r' \in \mathrm{CAR} \mid \pi_1(r') \in I \wedge \pi_2(r') \in J\}$. $R(I, J)$ restricts the first and second component of pairs and triples in CAR.

Definition 2. For $r \in \mathrm{CAR}$ define $\mathrm{ETC}(r) \in \mathbf{P}(\mathrm{CAR})$ as

$$\mathrm{ETC}(r) = \{r' \in \mathrm{CAR} \mid |r'| = |r| \wedge |r| = 3 \Rightarrow \pi_3(r') = \pi_3(r)\}.$$

Equal Third Component of r selects pairs r' if r is a pair (and hence contains no third component) and otherwise triples r' with the same third component as r.

Definition 3. For $B \in \mathbf{B}(\mathrm{CAR})$, $I \in \mathbf{P}(\mathbb{N})$ and $i \in \mathbb{N}$ we define $B[I \to i] \in \mathbf{B}(\mathrm{CAR})$ as

$$B[I \to i](r) = \begin{cases} 0, & \text{if } \pi_1(r) \in I \setminus \{i\} \vee \pi_2(r) \in I \setminus \{i\} \\ B(r) + \sum_{r' \in \mathrm{ETC}(r) \cap R(I \setminus \{i\}, \{\pi_2(r)\})} B(r'), & \\ \quad \text{if } \pi_1(r) = i \wedge \pi_2(r) \notin I \cup \{i\} & \\ B(r) + \sum_{r' \in \mathrm{ETC}(r) \cap R(\{\pi_1(r)\}, I \setminus \{i\})} B(r'), & \\ \quad \text{if } \pi_1(r) \notin I \cup \{i\} \wedge \pi_2(r) = i & \\ B(r) + \sum_{r' \in \mathrm{ETC}(r) \cap (R(\{i\}, I \setminus \{i\}) \cup R(I \setminus \{i\}, \{i\}) \cup R(I \setminus \{i\}, I \setminus \{i\}))} B(r'), & \\ \quad \text{if } \pi_1(r) = \pi_2(r) = i & \\ B(r), & \text{otherwise.} \end{cases}$$

When substituting i for the elements of I in B, the components in the records that get changed are the elements of $I \setminus \{i\}$: these components are replaced by i. With this in mind, the second line is concerned with records before the substitution of the form $\langle j, k \rangle$ or $\langle j, k, v \rangle$, the third line with $\langle k, j \rangle$ or $\langle k, j, v \rangle$, and the fourth line with $\langle i, j \rangle$ or $\langle i, j, v \rangle$ or $\langle j, i \rangle$ or $\langle j, i, v \rangle$ or $\langle j, m \rangle$ or $\langle j, m, v \rangle$, where $j, m \in I \setminus \{i\}$ and $k \notin I \cup \{i\}$.

When r, the record after substitution, has a third component only records r' before the substitution should be considered above that have a third component with the same value. This is taken care of in the equation by ETC.

ACKNOWLEDGMENTS

We are indebted to the second author, who, during a four-month visit to the universities of Nijmegen and Utrecht in the fall of 1983, started this research by writing (Shyamasundar and de Roever, 1983). Our thanks goes to Amir Pnueli, who assisted at several occasions in correcting and improving previous versions of this paper, and to the two referees for their comments and improvements. The Netherlands Organization for the Advancement of Pure Research (ZWO) is thanked for support of three of the authors. Finally, we thank Mijem Tosendjojo and Edmé van Thiel-Niekoop for their assistance with the typing of this paper.

RECEIVED October 4, 1985; ACCEPTED December 12, 1987

REFERENCES

ADA (1983), "The Programming Language Ada Reference Manual," Lecture Notes in Comput. Sci. Vol. 155, Springer-Verlag, New York/Berlin.

BERRY, G., AND COSSERAT, L. (1985), The ESTEREL synchronous programming language and its mathematical semantics, *in* "Seminar on Concurrency, July 1984, Carnegie-Mellon

University," Lecture Notes in Comput. Sci. Vol. 197, pp. 389–448, Springer-Verlag, New York/Berlin.

BERNSTEIN, A., AND HARTER, P. K., JR. (1981), Proving real-time properties of programs with temporal logic, *in* "8th ACM Symp. on Operating Systems Principles," pp. 1–11.

BROOKES, S. D., HOARE, C. A. R., AND ROSCOE, A. W. (1984), A theory of communicating sequential processes, *J. Assoc. Comput. Mach.* **31**, 560–599.

BARRINGER, H., KUIPER, R., AND PNUELI, A. (1984), Now you may compose temporal logic specifications, *in* "16th ACM Symp. Theory of Comput.," pp. 51–63.

BRANQUART, P., LOUIS, G., AND WODON, P. (1982), An analytical description of CHILL, the CCITT high level language VI, Lecture Notes in Comput. Sci. Vol. **128**, Springer-Verlag, New York/Berlin.

BJØRNER, D., AND OEST, O. N. (Eds.) (1980), "Towards a Formal Description of Ada," Lecture Notes in Comput. Sci. Vol. 98, Springer-Verlag, New York/Berlin.

BOUCHER, A. (1986), D. Phil. thesis, Department of Computer Science, University of Oxford.

CACM (1984), A case study: The space shuttle software system, *Comm. ACM* **27**, No. 9.

CAMERINI, J. (1982), "Sémantique Mathématique de Primitives Temps Reel," Thèse de 3ième cycle, IMA, Université de Nice.

DIJKSTRA, E. W. (1959), "Communication with an Automatic Computer," Ph.D. thesis, Mathematical Centre, Amsterdam.

FRANCEZ, N., LEHMANN, D., AND PNUELI, A. (1984), A linear-history semantics for languages for distributed programming, *Theoret. Comput. Sci.* **32**, 25–46.

GERTH, R., AND BOUCHER, A. (1987), A timed failures model for extended communicating processes, *in* "14th Int. Colloq. Automata. Lang. Programming," Lecture Notes in Comput. Sci. Vol. 267, pp. 95–114, Springer-Verlag, New York/Berlin.

GERTH, R. (1985), A maximal parallelism semantics for Occam, notes.

HAREL, D. (1987), Statecharts: A visual formalism for complex systems, *Sci. Comput. Programming* **8**, 231–274.

HUIZING, C., GERTH, R., AND DE ROEVER, W. P. (1987), Full abstraction of a real-time denotational semantics for an OCCAM-like language, *in* "14th ACM Principles of Programming Lang." pp. 223–237.

HEHNER, E. C. R., AND HOARE, C. A. R., (1983), A more complete model of communicating processes, *Theoret. Comput. Sci.* **26**, 105–120.

HALPERN, J. Y., MEGIDDO, N., AND MUNSHI, A. A. (1985), "Optimal Precision in the Presence of Uncertainty," IBM Research Lab., San Jose.

HOARE, C. A. R. (1978), Communicating sequential processes, *Comm. ACM* **21**, No. 8.

HOOMAN, J. (1987), A compositional proof theory for real-time distributed message passing, *in* "Proceedings of PARLE, Vol. II." Lecture Notes in Comput. Sci. Vol. 259, pp. 315–332, Springer-Verlag, New York/Berlin.

HOOMAN, J. (1988), A compositional proof-system for an OCCAM-like real-time language, Computing Science Notes 87/14, Department of Mathematics and Computing Science, Eindhoven University of Technology.

JONES, G. (1982), D.Phil. thesis, Oxford, unpublished.

KOYMANS, R. (1984), Denotational semantics for real-time programming constructs in concurrent languages, notes.

KOYMANS, R., VYTOPIL, J., AND DE ROEVER, W. P. (1983), Real-time programming and asynchronous message passing, *in* "2nd ACM Principles of Distrib. Comput.," pp. 187–197.

MISRA, J., AND CHANDY, K. M. (1981), Proofs of networks of processes, *IEEE Trans. Software Engrg.* **SE-7**, No. 4, 417–426.

MILNER, R. (1973), An approach to the semantics of parallel programs, *in* "Proceedings, Convegno di Informatica Teorica, Pisa."

MILNER, R. (1983), Calculi for synchrony and asynchrony, *Theoret. Comput. Sci.* **25**, 267–310.

OCCAM (1984), "The Occam Language Reference Manual," Prentice–Hall, Englewood Cliffs, NJ.

SALWICKI, A., AND MÜLDNER, T. (1981), On the algorithmic properties of concurrent programs, Lecture Notes in Comput. Sci. Vol. 125, pp. 169–197, Springer-Verlag, New York/Berlin.

SHYAMASUNDAR, R. K., AND DE ROEVER, W. P. (1983), Semantics of real-time Ada, notes.

ZWIERS, J., DE ROEVER, W. P., AND VAN EMDE BOAS, P. (1985), Compositionality and concurrent networks: Soundness and completeness of a proofsystem, *in* "12th Int. Colloq. Automata. Lang. Programming," Lecture Notes in Comput. Sci., Vol. 194, pp. 509–519, Springer-Verlag, New York/Berlin.

ZWIERS, J. (1988), "Compositionality, Concurrency and Partial Correctness: Proof Theories for Networks of Processes, and Their Connection," Ph.D. thesis, Eindhoven University of Technology.

ZIJLSTRA, E. (1984), "Real-Time Semantics," Master thesis, University of Amsterdam.