A calculus for team automata*

Maurice H. ter Beek[1], **Fabio Gadducci**[2], **Dirk Janssens**[3]

[1]Istituto di Scienza e Tecnologie dell'Informazione, CNR
via G. Moruzzi 1, 56124 Pisa, Italy

[2]Dipartimento di Informatica, Università di Pisa
via Buonarroti 2, 56125 Pisa, Italy

[3]Department of Mathematics and Computer Science, University of Antwerp
Middelheimlaan 1, 2020 Antwerpen, Belgium

m.terbeek@isti.cnr.it, gadducci@di.unipi.it, dirk.janssens@ua.ac.be

Abstract. *Team automata are a formalism for the component-based specification of reactive, distributed systems. Their main feature is a flexible technique for specifying coordination patterns among systems, thus extending I/O automata. Furthermore, for some patterns the language recognized by a team automaton can be specified via those languages recognized by its components.*
We introduce a process calculus tailored over team automata. Each automaton is described by a process, and such that its associated (fragment of a) labeled transition system is bisimilar to the original automaton. The mapping is furthermore denotational, since the operators defined on processes are in a bijective correspondence with a chosen family of coordination patterns and that correspondence is preserved by the mapping.
We thus extend to team automata a few classical results on I/O automata and their representation by process calculi. Moreover, besides providing a language for expressing team automata, we widen the family of coordination patterns for which an equational characterization of the language associated to a composite automaton can be provided. The latter result is obtained by providing a set of axioms, in ACP-style, for capturing bisimilarity in our calculus.

1. Introduction

Team automata have originally been introduced in the context of Computer Supported Cooperative Work (CSCW for short) to formalize the conceptual and architectural levels of groupware systems [ter Beek et al. 2003, Ellis 1997, Kleijn 2003]. As shown in [ter Beek and Kleijn 2005], they represent an extension of classical I/O automata [Lynch 1996, Lynch and Tuttle 1989], and since their introduction they have proved their usefulness also in various application fields [ter Beek 2003, ter Beek et al. 2005, ter Beek et al. 2006]. Team automata form a mathematical framework enabling one to capture notions like communication, coordination and cooperation in reactive, distributed systems. The model allows one to separately specify the components of a system, to describe their interactions and to reuse the system as a component of

*Work partly supported by the EU projects HPRN-CT-2002-00275 SEGRAVIS (*Syntactic and Semantic Integration of Visual Modelling Techniques*) and IST-3-016004-IP-09 SENSORIA (*Software Engineering for Service-Oriented Overlay Computers*).

a higher-level team automaton, thus supporting a modular approach to system design. Its main feature is a flexible technique for specifying coordination patterns among distributed systems, extending classical I/O automata.

During the stepwise development of a system, it is desirable to have the possibility to decompose an abstract high-level specification of a large, complex design into a more concrete low-level specification. Of course, in order to guarantee that the decomposition is correct, it is necessary to prove that the chosen model is compositional, i.e., that the specification of a large system can be obtained from specifications of its components [Jonsson 1994]. Unfortunately, as we show (see Proposition 1), for some of the coordination patterns employed so far it is not possible to capture the behavior of a (finite!) team automaton (intended as the language it recognizes) in terms of its components, by resorting only to set-theoretic operations on languages.

In order to overcome this difficulty, we introduce a calculus for team automata. Our proposal recalls those calculi that have been defined for (probabilistic) I/O automata [De Nicola and Segala 1995, Stark et al. 2003, Vaandrager 1991], and the aim is to transfer the technology involving the equational characterization of behavioral equivalences on processes to team automata, in order to obtain a characterization of the relevant coordination patterns. The main idea underlying process algebras like the ACP framework [Bergstra and Klop 1984] and Hoare's CSP [Hoare 1985] is to use a set of operators, each one representing an architectural feature, for an inductive presentation of a complex system. Our calculus for team automata is essentially an enrichment of CSP, and its behavioral semantics is axiomatized by suitably adapted operators from the ACP framework. Each automaton is described by a process, in such a way that its associated (fragment of a) labeled transition system is behaviorally equivalent (namely, *bisimilar* [Milner 1980]) to the original automaton. Furthermore, the mapping is denotational, since the operators on processes are in a bijective correspondence with a chosen family of coordination patterns, and that correspondence is preserved by the mapping.

One of our results is thus the extension to team automata of some classical results on I/O automata and their representation by process calculi. Another result concerns the compositionality of team automata. In [ter Beek 2003, ter Beek and Kleijn 2003] it was shown that certain team automata that are defined by a coordination pattern are compositional, in the sense that their languages can be obtained from the languages of their constituting automata. Besides proving that this characterization does not hold for all coordination patterns devised so far (even in the presence of acyclic automata: See the already mentioned Proposition 1), we use our calculus to provide some preliminary results on how to nevertheless obtain the language of a team automaton defined by a coordination pattern directly from its components. Hence, a compositionality result does exist, even if the manipulation of the languages of the components does not suffice. By providing a set of axioms, in ACP-style, to capture bisimilarity in our calculus, we thus enlarge the family of coordination patterns for which an equational characterization of the language associated to a team automaton can be provided. This axiomatization is sound and complete for finite processes only (i.e., equivalently, for acyclic automata) as is typical for all the calculi for automata that we are aware of (compare, e.g., the situation for (probabilistic) I/O automata, as reported in [De Nicola and Segala 1995, Stark et al. 2003]). The search for a set of axioms for possibly recursive processes is currently under development.

The paper is organized as follows. Section 2 recalles the main definitions concerning team automata. Then, the syntax and the operational semantics of our calculus for team automata, as well as an equational theory for bisimulation over finite processes, are given in Section 3. Section 4 presents an encoding from team automata to processes that preserves bisimulation equivalence, while Section 5 offers a characterization — via a suitable axiomatization — for language equivalence, thus partly solving (since the encoding preserves the composition patterns on automata) our modularity issues. Finally, Section 6 concludes the paper, hinting at possible future work.

2. Team automata

Roughly speaking, a team automaton consists of component automata — ordinary automata without final states and with a distinction of their sets of actions into input, output and internal actions — combined in such a way that they can perform shared actions. During each clock tick the components within a team can simultaneously participate in one instantaneous action (i.e., synchronize on this action) or remain idle. Component automata can thus be combined in a loose or more tight fashion, depending on which actions are to be synchronized and when.

We now fix some notations and terminology used in this article, after which we introduce team automata. However, we slightly adapt the usual definition of team automata [ter Beek et al. 2003]. First, we assume each automaton to have a unique initial state. This is of course not a real limitation, but it will ease some of the constructions below. Second, we discard the usual distinction between input, output and internal actions in component and team automata. In [De Nicola and Segala 1995, Stark et al. 2003, Vaandrager 1991] the distinction of the set of actions of I/O automata into input, output and internal actions is taken into account. For team automata, however, this distinction is much less important since — contrary to I/O automata — team automata are not required to be input enabling and synchronizations between output actions are not prohibited [ter Beek et al. 2003, ter Beek and Kleijn 2005]. Hence in team automata the consideration of input and output actions does not have any syntactic significance. As a result, taking these actions into account would not affect our calculus. Moreover, it would not be very difficult to extend our calculus in order to deal with internal actions.

For convenience we sometimes denote the set $\{1, \ldots, n\}$ by $[n]$; thus $[0] = \varnothing$. The (cartesian) product of sets V_i, with $i \in [n]$, is denoted either by $\prod_{i\in[n]} V_i$ or by $V_1 \times \cdots \times V_n$. For $j \in [n]$, $\text{proj}_j : \prod_{i\in[n]} V_i \to V_j$ is defined by $\text{proj}_j((a_1, \ldots, a_n)) = a_j$. The set difference of sets V and W is denoted by $V \setminus W$. For a finite set V, its cardinality is denoted by $\#V$.

Let Γ and Σ be sets of symbols. The morphism $\text{pres}_{\Gamma,\Sigma} : \Gamma^* \to \Sigma^*$, defined by $\text{pres}_{\Gamma,\Sigma}(a) = a$ if $a \in \Sigma$ and $\text{pres}_{\Gamma,\Sigma}(a) = \lambda$ otherwise, preserves the symbols from Σ and erases all other symbols. We discard Γ when no confusion can arise.

Let $f : A \to A'$ and $g : B \to B'$ be two functions. Then $f \times g : A \times B \to A' \times B'$ is defined as $(f \times g)(a, b) = (f(a), g(b))$. We use $f^{[2]}$ as shorthand for $f \times f$.

Definition 1 *A* labeled transition system *(*LTS *for short) is a triple $\mathcal{A} = (Q, \Sigma, \delta)$, with a set Q of* states*, a set Σ of actions (satisfying $Q \cap \Sigma = \varnothing$) and a set $\delta \subseteq Q \times \Sigma \times Q$ of* labeled transitions*.*

The set δ_a of *a-transitions* of $\mathcal{A}$ is defined as $\delta_a = \{ (q, q') \mid (q, a, q') \in \delta \}$ and an a-transition $(q, a, q') \in \delta$ may also be written as $q \xrightarrow{a} q'$. Action a is said to be *enabled in $\mathcal{A}$ at* state $q \in Q$, denoted by a en$_{\mathcal{A}}$ q, if there exists $q' \in Q$ such that $(q, q') \in \delta_a$. An a-transition $(q, q) \in \delta_a$ is called a *loop* (on a).

Definition 2 *A* (component) automaton $\mathcal{C}$ *is a* finite, rooted LTS*, i.e., a quadruple* (Q, Σ, δ, q_0)*, where* (Q, Σ, δ) *is an* LTS *with finite* Q *and* Σ *and an* initial state $q_0 \in Q$.

The set $\mathbb{C}(\mathcal{C})$ *of* computations *of* $\mathcal{C}$ *is defined as* $\mathbb{C}(\mathcal{C}) = \{ q_0 a_1 q_1 a_2 \cdots a_n q_n \mid n \geq 0$ *and* $(q_{i-1}, a_i, q_i) \in \delta$ *for all* $i \in [n] \}$.

The language $\mathbb{L}(\mathcal{C})$ *of* $\mathcal{C}$ *is defined as* $\mathbb{L}(\mathcal{C}) = pres_{\Sigma}(\mathbb{C}(\mathcal{C}))$.

In the sequel, we let $\mathcal{S} = \{ \mathcal{C}_i \mid i \in [n] \}$ be an arbitrary but fixed set of automata, with $n \geq 0$ and each $\mathcal{C}_i$ specified as $\mathcal{C}_i = (Q_i, \Sigma_i, \delta_i, q_{0i})$, and we let $\Sigma = \bigcup_{i \in [n]} \Sigma_i$.

A team automaton over $\mathcal{S}$ has the cartesian product of the state spaces of its components as its state space and its actions are those of its components. Its transition relation, however, is based on but not fixed by those of its components: The transition relation of a team automaton over $\mathcal{S}$ is defined by choosing certain synchronizations of actions of its components, while excluding others.

Definition 3 *Let* $a \in \Sigma$*. The set* $\Delta_a(\mathcal{S})$ *of* synchronizations *of* a *is defined as* $\Delta_a(\mathcal{S}) = \{ (q, q') \in \prod_{i \in [n]} Q_i \times \prod_{i \in [n]} Q_i \mid [\exists\, j \in [n] : proj_j{}^{[2]}(q, q') \in \delta_{j,a}] \wedge [\forall\, i \in [n] : [proj_i{}^{[2]}(q, q') \in \delta_{i,a}] \vee [proj_i(q) = proj_i(q')]\,] \}$.

The set $\Delta_a(\mathcal{S})$ contains all possible combinations of a-transitions of the components constituting $\mathcal{S}$, with all non-participating components remaining idle. It is explicitly required that in every synchronization at least one component participates. The state change of a team automaton over $\mathcal{S}$ is thus defined by the local state changes of the components constituting $\mathcal{S}$ that participate in the action of the team being executed. Hence, when defining a team automaton over $\mathcal{S}$, a specific subset of $\Delta_a(\mathcal{S})$ must be chosen for each action a. This defines an explicit communication pattern between those components constituting the team.

Definition 4 *A* team automaton *over* $\mathcal{S}$ *is a quadruple* $\mathcal{T} = (Q, \Sigma, \delta, q_0)$*, with* $Q = \prod_{i \in [n]} Q_i$*,* $\Sigma = \bigcup_{i \in [n]} \Sigma_i$*,* $\delta \subseteq Q \times \Sigma \times Q$ *such that* $\delta_a = \{ (q, q') \mid (q, a, q') \in \delta \} \subseteq \Delta_a(\mathcal{S})$ *for all* $a \in \Sigma$ *and* $q_0 = \prod_{i \in [n]} q_{0i}$.

In [ter Beek et al. 2003] several coordination patterns for the synchronizations of a team automaton were defined, each leading to a uniquely defined team automaton. These patterns fix the synchronizations of a team by defining — per action a — certain conditions on the a-transitions to be chosen from $\Delta_a(\mathcal{S})$, thus determining a unique subset of $\Delta_a(\mathcal{S})$ as the set of a-transitions of the team. Once such subsets have been chosen for all actions, the team automaton they define is unique.

Definition 5 *Let* $\Gamma \subseteq \Sigma$*,* $\mathcal{R}_a(\mathcal{S}) \subseteq \Delta_a(\mathcal{S})$ *for all* $a \in \Gamma$ *and* $\mathcal{R}_\Gamma = \{ \mathcal{R}_a(\mathcal{S}) \mid a \in \Gamma \}$*. Then* $\mathcal{T} = (Q, \Sigma, \delta, q_0)$ *is the* $\mathcal{R}_\Sigma$-team automaton *over* $\mathcal{S}$ *if* $\delta_a = \mathcal{R}_a(\mathcal{S})$ *for all* $a \in \Sigma$.

In this notation we usually discard Σ if no confusion can arise. Here we consider three coordination patterns, based on those actions of $\mathcal{T}$ that are *free*, *ai* or *si*. An action a is *free* in $\mathcal{T}$ if none of its a-transitions is brought about by a synchronization of a by two

or more components from $\mathcal{S}$, action a is *action-indispensable* (*ai* for short) in $\mathcal{T}$ if all its a-transitions are brought about by a synchronization of all components from $\mathcal{S}$ sharing a and action a is *state-indispensable* (*si* for short) in $\mathcal{T}$ if all its a-transitions are brought about by a synchronization of all components from $\mathcal{S}$ in which a is currently enabled.

Definition 6 *Let $a \in \Sigma$. Then we define the sets*

- $\mathcal{R}_a^{no}(\mathcal{S}) = \Delta_a(\mathcal{S})$;
- $\mathcal{R}_a^{free}(\mathcal{S}) = \{\, (q,q') \in \Delta_a(\mathcal{S}) \mid \#\{\, i \in [n] \mid a \in \Sigma_i \wedge \mathit{proj}_i^{[2]}(q,q') \in \delta_{i,a} \,\} = 1 \,\}$;
- $\mathcal{R}_a^{ai}(\mathcal{S}) = \{\, (q,q') \in \Delta_a(\mathcal{S}) \mid \forall\, i \in [n] : [a \in \Sigma_i \Rightarrow \mathit{proj}_i^{[2]}(q,q') \in \delta_{i,a}] \,\}$;
- $\mathcal{R}_a^{si}(\mathcal{S}) = \{\, (q,q') \in \Delta_a(\mathcal{S}) \mid \forall\, i \in [n] : [[a \in \Sigma_i \wedge a\ \mathit{en}_{\mathcal{A}_i}\ \mathit{proj}_i(q)] \Rightarrow \mathit{proj}_i^{[2]}(q,q') \in \delta_{i,a}] \,\}$.

Each of these subsets of $\Delta_a(\mathcal{S})$ thus defines, for a given action $a \in \Sigma$, *all* transitions from $\Delta_a(\mathcal{S})$ that satisfy a certain type of synchronization. In the case of *no* constraints, this means that all a-transitions are allowed since nothing is required, and hence no transition is forbidden. In the other three cases, *all and only those* a-transitions are included that respect the specified property of a.

Before presenting an example to illustrate the notions defined so far, we define shorthand notations for three specific types of team automata that we will use in the sequel. Let $n = 2$ (i.e., we consider $\mathcal{S} = \{\mathcal{C}_1, \mathcal{C}_2\}$) and let $\Gamma \subseteq \Sigma$. Then

- $\mathcal{C}_1 \,|||_\Gamma^f\, \mathcal{C}_2$ defines the $\mathcal{R}_{\Sigma\setminus\Gamma}^{no} \cup \mathcal{R}_\Gamma^{free}$-team automaton over $\mathcal{S}$;
- $\mathcal{C}_1 \,|||_\Gamma^{ai}\, \mathcal{C}_2$ defines the $\mathcal{R}_{\Sigma\setminus\Gamma}^{no} \cup \mathcal{R}_\Gamma^{ai}$-team automaton over $\mathcal{S}$;
- $\mathcal{C}_1 \,|||_\Gamma^{si}\, \mathcal{C}_2$ defines the $\mathcal{R}_{\Sigma\setminus\Gamma}^{no} \cup \mathcal{R}_\Gamma^{si}$-team automaton over $\mathcal{S}$.

Example 1 *Consider the two component automata $\mathcal{C}_1 = (\{p,p'\}, \{b\}, \{(p,b,p')\}, p)$ and $\mathcal{C}_2 = (\{q,q'\}, \{a,b\}, \{(q,b,q),(q,a,q')\}, q)$. They are depicted in Figure 1.*

Figure 1. From left to right: component automata $\mathcal{C}_1$ and $\mathcal{C}_2$.

In Figure 2 we have depicted $\mathcal{C}_1 \,|||_{\{a,b\}}^f\, \mathcal{C}_2$, $\mathcal{C}_1 \,|||_{\{b\}}^{ai}\, \mathcal{C}_2$ and $\mathcal{C}_1 \,|||_{\{b\}}^{si}\, \mathcal{C}_2$. Note that $\mathcal{C}_1 \,|||_{\{a,b\}}^f\, \mathcal{C}_2$ has no b-transition from (p,q) to (p',q). In fact, this team automaton is different from the $\mathcal{R}_{\{a,b\}}^{no}$-team automaton over $\{\mathcal{C}_1, \mathcal{C}_2\}$ due to the loop on b in $\mathcal{C}_2$ (more about this in Section 3).

A team automaton over $\mathcal{S}$ is said to satisfy *compositionality* if its behavior (i.e., its language) can be described in terms of that of its constituting component automata: There exists a set-theoretic operation that when applied to the languages of the automata in $\mathcal{S}$, the language of a particular team over $\mathcal{S}$ results. In [ter Beek 2003, ter Beek and Kleijn 2003] it was shown that the construction of team automata according to certain patterns of synchronization, e.g., the ones leading to $\mathcal{R}^{free}$- and $\mathcal{R}^{ai}$-team automata, guarantees compositionality. In [ter Beek 2003] it is moreover claimed that a similar result for the case of $\mathcal{R}^{si}$-team automata "seems impossible due to the simple fact that the behavior of component automata is stripped from all state information". Here we prove this statement.

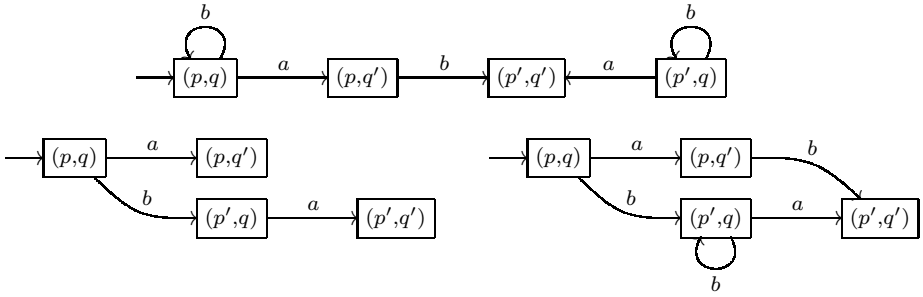

Figure 2. Clockwise from top: team automata $\mathcal{C}_1|||^{f}_{\{a,b\}}\mathcal{C}_2$, $\mathcal{C}_1|||^{si}_{\{b\}}\mathcal{C}_2$ and $\mathcal{C}_1|||^{ai}_{\{b\}}\mathcal{C}_2$.

Proposition 1 *Let $\mathcal{C}_1$ and $\mathcal{C}_2$ be two component automata. Then there exists no set-theoretic operation $|||$ on languages such that $\mathbb{L}(\mathcal{C}_1 \,|||^{si}_{\Sigma}\, \mathcal{C}_2) = \mathbb{L}(\mathcal{C}_1) \,|||\, \mathbb{L}(\mathcal{C}_2)$.*

The proof is by counterexample. Consider the component automata in Figure 3. Then $\mathbb{L}(\mathcal{D}_2) = \mathbb{L}(\mathcal{D}_3)$, while $\mathbb{L}(\mathcal{D}_1 \,|||^{si}_{\Sigma}\, \mathcal{D}_2) = \mathbb{L}(\mathcal{D}_1 \,|||^{si}_{\Sigma}\, \mathcal{D}_3) \cup \{abc\}$.

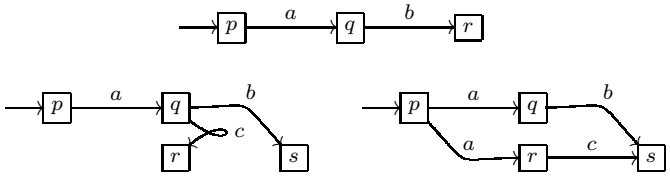

Figure 3. Clockwise from top: component automata $\mathcal{D}_1$, $\mathcal{D}_2$ and $\mathcal{D}_3$.

In Section 5 we show that the calculus for team automata that we are going to introduce does provide a recipe to obtain the language of an acyclic $\mathcal{R}^{si}$-team automaton 'directly' (i.e., without actually constructing the team automaton) from its constituting component automata: Translate the component automata to processes, perform the so-called eager parallel composition operation defined below, derive the normal form of the resulting process term and its associated regular language is the desired language.

3. A CSP-like process calculus

In this section we introduce a simple process calculus, essentially an enrichment of Hoare's CSP [Hoare 1985], and then present the associated operational semantics.

3.1. Syntax and operational semantics

We assume countable sets of *actions* A, ranged over by $a, b, \ldots$, and *agent variables* X, ranged over by $x, y, \ldots$, with $\wp_f(A)$ — the finite subsets of A — ranged over by L. Terms are built from actions and variables according to the syntax

$$\begin{array}{rcl} M & ::= & \mathsf{nil} \mid a.x \mid a.P \mid a.M \mid M + M \mid rec_x.M \\ P & ::= & M_c \mid P \parallel^{f}_{L} P \mid P \parallel^{ai}_{L} P \mid P \parallel^{si}_{L} P \end{array}$$

As usual, a variable x is *free* if it does not occur inside the scope of a rec_x operator. The set of *(sequential) agents* is ranged over by $M, N, \ldots$, and for its subsets of *closed* agents the subscript $_c$ is added. The set of *processes* is denoted by $\mathcal{P}$ and ranged over by $P, Q, \ldots$, and a process is *finite* if it contains no occurrence of a recursion operator.

The constant nil represents the terminated process. The *action prefix* $a.P$ can perform an atomic action a and then evolve to P. Summation $+$ denotes *non-deterministic choice*: $M + N$ behaves either as M or as N, the choice being triggered by the execution of an action. The intended meaning of the *recursion* operator $rec_x.M$ is the process defined by the equation $x = M$, with the further restriction implicitly ensured by the syntax, namely that only closed terms may be inserted into a parallel composition operator: This assumption corresponds to what are usually called *size-bounded* processes, and it is formalized by Proposition 2 below.

There are three different notions of parallel composition. Basically, $P \parallel_L^{ai} Q$ means that processes P and Q must evolve synchronously with respect to all actions in L, while they may evolve independently of each other with respect to actions $a \notin L$, i.e., the actions in L are synchronized according to the *ai* type of synchronization. Similarly for its *eager* version: Also in $P \parallel_L^{si} Q$ both processes must synchronize on the actions in L, but now a process may in any case evolve with any action that is not offered at the moment by the other process, i.e., the actions in L are synchronized according to the *si* type of synchronization. Finally, in $P \parallel_L^{f} Q$ the two processes may synchronize on actions $a \notin L$, but both processes must evolve independently of each other for all actions in L, in which case a further restriction is imposed in case one of the processes may loop: In order to faithfully mimic the *free* type of synchronization for all actions in L, a process may independently evolve with an action $a \in L$ only if the other process cannot evolve with a loop on a. This condition seems peculiar in the context of process calculi, but it is a consequence of the lack of explicit information on loops in team automata, i.e., in general it is impossible to distinguish whether or not a component with a loop on a in its current local state participates in the synchronization of the team on a. In [ter Beek et al. 2003] this led to the adoption of the maximal interpretation of the components' participation: Given a team transition (q, a, q') it is assumed that the jth component participates in this transition by executing $(\text{proj}_j(q), a, \text{proj}_j(q'))$ whenever $\text{proj}^{[2]}(q, q') \in \delta_{j,a}$, whereas otherwise no transition takes place in the jth component (see Example 1).

The operational semantics of this calculus is described by the LTS $\mathcal{T} = (\mathcal{P}, A, \rightarrow)$, where $\rightarrow \subseteq \mathcal{P} \times A \times \mathcal{P}$ is defined in the so-called SOS style [Plotkin 1981] as the least relation that satisfies the set of axioms and inference rules of Table 1 (where we omitted the symmetric rules for the choice operator and for the three parallel composition operators). Note also the negative premises occurring in the last two rules, namely *asyn*$_L^f$ and *asyn*$_L^{si}$: $Q \not\xrightarrow{a}$ means that from Q there is no outgoing transition labeled with a and $Q \not\xrightarrow{a} Q$ means that from Q there is no outgoing transition labeled with a that results in a cycle. Due to the restricted structure of the processes, and since the inference rules increase the size of a process, the least transition relation is well-defined. The semantics of a process $P \in \mathcal{P}$, denoted by *LTS*(P), is defined as the rooted LTS *LTS*$(P) = (\mathcal{P}, A, \rightarrow, P)$.

Example 2 *Consider the simple sequential agents* $M = b.\mathsf{nil}$ *and* $N = rec_x(b.x + a.\mathsf{nil})$. *Their associated rooted* LTS*'s are depicted in Figure 4.*

Figure 4. From left to right: *LTS*(M)**,** $M = b.\mathsf{nil}$**, and** *LTS*(N)**,** $N = rec_x(b.x + a.\mathsf{nil})$**.**

Table 1. The operational semantics for $\mathcal{P}$.

$$\textit{act}: \ \frac{-}{a.P \xrightarrow{a} P} \qquad \textit{sum}: \ \frac{M \xrightarrow{a} M'}{M + N \xrightarrow{a} M'} \qquad \textit{rec}: \ \frac{M[rec_x.M/x] \xrightarrow{a} N}{rec_x.M \xrightarrow{a} N}$$

$$\textit{par}^f: \ \frac{P \xrightarrow{a} P',\ Q \xrightarrow{a} Q'}{P \parallel^f_L Q \xrightarrow{a} P' \parallel^f_L Q'} \ a \notin L \qquad \textit{asyn}^f: \ \frac{P \xrightarrow{a} P'}{P \parallel^f_L Q \xrightarrow{a} P' \parallel^f_L Q} \ a \notin L$$

$$\textit{par}^{ai}: \ \frac{P \xrightarrow{a} P',\ Q \xrightarrow{a} Q'}{P \parallel^{ai}_L Q \xrightarrow{a} P' \parallel^{ai}_L Q'} \qquad \textit{asyn}^{ai}: \ \frac{P \xrightarrow{a} P'}{P \parallel^{ai}_L Q \xrightarrow{a} P' \parallel^{ai}_L Q} \ a \notin L$$

$$\textit{par}^{si}: \ \frac{P \xrightarrow{a} P',\ Q \xrightarrow{a} Q'}{P \parallel^{si}_L Q \xrightarrow{a} P' \parallel^{si}_L Q'} \qquad \textit{asyn}^{si}: \ \frac{P \xrightarrow{a} P'}{P \parallel^{si}_L Q \xrightarrow{a} P' \parallel^{si}_L Q} \ a \notin L$$

$$\textit{asyn}^f_L: \ \frac{P \xrightarrow{a} P',\ Q \not\xrightarrow{a} Q}{P \parallel^f_L Q \xrightarrow{a} P' \parallel^f_L Q} \ a \in L \qquad \textit{asyn}^{si}_L: \ \frac{P \xrightarrow{a} P',\ Q \not\xrightarrow{a}}{P \parallel^{si}_L Q \xrightarrow{a} P' \parallel^{si}_L Q} \ a \in L$$

Let $L = \{b\}$. Hence, no constraint is imposed on a. Then the LTS*'s corresponding to the application of the three parallel composition operators to M and N are depicted in Figure 5.*

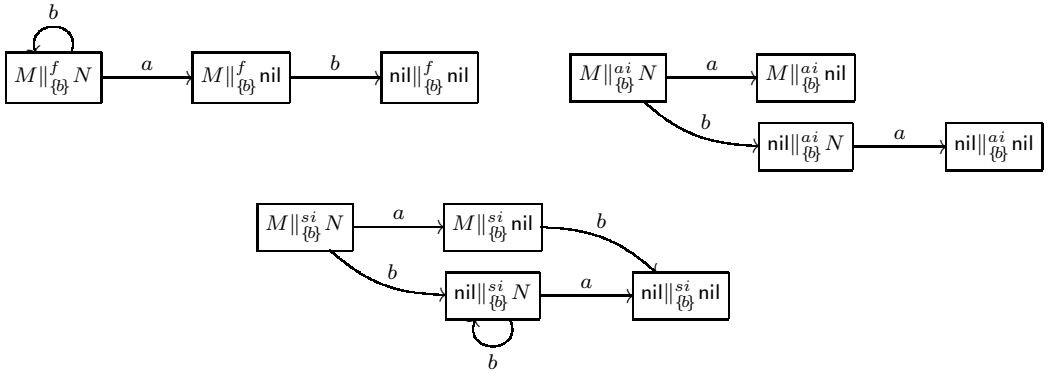

Figure 5. Clockwise from top: the LTS**'s for $M \parallel^f_{\{b\}} N$, $M \parallel^{ai}_{\{b\}} N$ and $M \parallel^{si}_{\{b\}} N$.**

The next section focuses on an equational presentation for *bisimulation* equivalence, equating those processes exhibiting the same (non-deterministic) operational behavior. The result below states a property of our operational semantics, making precise the previous remark on size-bounded processes.

Proposition 2 *Let P be a process. Then the rooted* LTS *LTS(P) is finite.*

In other words, no syntactic explosion of a process during its evolution may occur, because only closed terms may be inserted into parallel composition operators.

3.2. Axioms for bisimulation

The aim of this section is to introduce a finite equational theory for bisimulation, which will later form the basis for the characterization of the language associated to a process (hence, to an automaton). First we define the notion of bisimulation.

Definition 7 *Let $\mathcal{T} = (Q, \Sigma, \delta)$ be an* LTS. *A relation $R \subseteq Q \times Q$ is a* bisimulation *if, whenever $(p, q) \in R$, then for any $p', q' \in Q$ and any $a \in \Sigma$ holds*

1. *if $p \xrightarrow{a} p'$, then $q \xrightarrow{a} q'$ for some $q' \in Q$ such that $(p', q') \in R$;*
2. *if $q \xrightarrow{a} q'$, then $p \xrightarrow{a} p'$ for some $p' \in Q$ such that $(p', q') \in R$.*

Two states $q, q' \in Q$ are said to be bisimilar, *denoted by $q \simeq q'$, if there exists a bisimulation R such that $(q, q') \in R$. Two rooted* LTS*'s $\mathcal{T}_1 = (Q_1, \Sigma_1, \delta_1, q_1)$ and $\mathcal{T}_2 = (Q_2, \Sigma_2, \delta_2, q_2)$ are* bisimilar *if $q_1 \simeq q_2$. Two processes P and Q are* bisimilar *if LTS(P) and LTS(Q) are.*

It often occurs that bisimulation is not a congruence with respect to the operators of the calculus, whenever there are rules containing negative premises. In fact, two of the SOS rules of Table 1 have negative premises, and the set of rules of our calculus does not fit the general so-called *ntyft/ntyxt* format [Bol and Groote 1996].

The main problem is the $asyn_L^f$ rule, since it contains an explicit hypothesis on the target state of the negative premise. It is in fact easy to see that the process $P = rec_x.a.x$ and its unfolding $Q = a.rec_x.a.x$ are bisimilar, while $P \parallel^f_{\{a\}} R$ and $Q \parallel^f_{\{a\}} R$ are not, for $R = a.b.\mathsf{nil}$. However, the problem disappears whenever we restrict our attention to finite processes, since the negative premise of the rule is always void. Thus, in the remaining of the section we restrict our attention to finite processes only.

Our starting point for a finite equational theory for bisimulation is the solution routinely adopted in the ACP framework [Bergstra and Klop 1984], i.e., to use suitable auxiliary operators (usually $\between$ and $|$) to split the parallel composition operator ($\|$) into its possible behaviors: either an asynchronous evolution ($\between$) or a forced synchronization ($|$). For our calculus of finite processes this leads to the axioms concerning the choice and parallel composition operators reported in Tables 2 and 3, respectively. Concerning parallel composition, the lack of a superscript (either f, ai or si) means that the law holds for each of the three operators. Furthermore, given a process $P \in \mathcal{P}$, the predicate En(P) is defined as En$(P) = \{\, a \in A \mid \exists Q \in \mathcal{P} :\ P \xrightarrow{a} Q \,\}$.

Table 2. Axioms for the choice operator.

$$M + M = M \qquad\qquad M + N = N + M$$

$$(M + N) + O = M + (N + O) \qquad\qquad M + \mathsf{nil} = M$$

Proposition 3 *Let P, Q be finite processes. Then P and Q are bisimilar if and only if they are equated by the axioms of Tables 2 and 3.*

Since the set of SOS rules of our calculus of finite processes can be transformed into a set of so-called *smooth* GSOS rules, we could as well have used the general procedure described in [Aceto et al. 1994] to automatically generate a complete axiomatization for bisimulation. We however chose to provide a direct, intuitive set of axioms.

Note that the equations of Tables 2 and 3 can be oriented from left to right, so that they actually induce a rewriting system, modulo the so-called AC (associativity and commutativity) axioms for the choice operator $+$. So, two finite processes are bisimilar if they have the same (modulo AC) *normal form* (i.e., the process that is obtained after rewriting the original process according to the rewriting system until no further rewriting can be applied).

Table 3. Axioms for the parallel composition operators.

$$P \parallel_L Q = P \between_L Q + Q \between_L P + P \mid_L Q \qquad (P + Q) \mid_L R = P \mid_L R + Q \mid_L R$$

$$(P + Q) \between_L R = P \between_L R + Q \between_L R \qquad R \mid_L (P + Q) = R \mid_L P + R \mid_L Q$$

$$a.P \between_L^f Q = a.(P \parallel_L^f Q) \qquad \mathsf{nil} \between_L P = \mathsf{nil}$$

$$a.P \mid_L^f a.Q = \begin{cases} a.(P \parallel_L^f Q) & \text{if } a \notin L \\ \mathsf{nil} & \text{otherwise} \end{cases} \qquad \mathsf{nil} \mid_L P = \mathsf{nil} = P \mid_L \mathsf{nil}$$

$$a.P \between_L^{ai} Q = \begin{cases} a.(P \parallel_L^{ai} Q) & \text{if } a \notin L \\ \mathsf{nil} & \text{otherwise} \end{cases} \qquad a.P \mid_L b.Q = \mathsf{nil}$$

$$a.P \between_L^{si} Q = \begin{cases} a.(P \parallel_L^{si} Q) & \text{if } a \notin L \cap \mathrm{En}(Q) \\ \mathsf{nil} & \text{otherwise} \end{cases} \qquad a.P \mid_L^{\{ai,si\}} a.Q = a.(P \parallel_L^{\{ai,si\}} Q)$$

4. From automata to processes

The aim of this section is to present an encoding from automata to processes such that bisimulation equivalence is preserved. To this end, we now extend the usual definition of automata by assigning a specific set of states to be considered as entry points for the recursion operator.

Definition 8 *Let X be a set of state variables. Then an automaton over X is a pair $\langle \mathcal{A}, f \rangle$, where $\mathcal{A} = (Q, \Sigma, \delta, q_0)$ is an automaton and $f : X \to Q$ is an injective (possibly partial) function.*

So, for the rest of this section we assume that for each automaton a set of its states is uniquely labeled by an element in X.

It is now possible to define our encoding from automata to processes.

Definition 9 *Let $\langle \mathcal{A}, f \rangle$ be an automaton $\mathcal{A} = (Q, \Sigma, \delta, q_0)$ over $X_{\mathcal{A}}$. Then the algorithm obtained by repeatedly applying the three steps below inductively defines an* essentially unique *— up to the choice of variables — process Exp$(\langle \mathcal{A}, f \rangle)$.*

- *If q_0 has no outgoing transitions, then*

$$\mathit{Exp}(\langle \mathcal{A}, f \rangle) = \begin{cases} x & \text{if } f(x) = q_0, \text{ for some } x \in X_{\mathcal{A}}, \text{ and} \\ \mathsf{nil} & \text{otherwise;} \end{cases}$$

- *If q_0 has $n > 0$ outgoing transitions (q_0, a_i, q_i) and no incoming ones, then*

$$\mathit{Exp}(\langle \mathcal{A}, f \rangle) = \sum_{i \in \{1,\ldots,n\}} a_i.\mathit{Exp}(\langle \mathcal{A}_i, f \rangle)$$

  for automata $\mathcal{A}_i = (Q \setminus \{q_0\}, \Sigma, \delta \setminus \{\, (q_0, a, q) \mid a \in \Sigma, q \in Q \,\}, q_i)$ over $X_{\mathcal{A}}$;

- *If q_0 has $n > 0$ outgoing transitions (q_0, a_i, q_i) and some incoming ones, then*

$$\mathit{Exp}(\langle \mathcal{A}, f \rangle) = rec_x.\left(\sum_{i \in \{1,\ldots,n\}} a_i.\mathit{Exp}(\langle \mathcal{A}_i, g \rangle) \right)$$

  for a new variable x, automata $\mathcal{A}_i = (Q, \Sigma, \delta \setminus \{\, (q_0, a, q) \mid a \in \Sigma, q \in Q \,\}, q_i)$ over $X_{\mathcal{A}} \cup \{x\}$ and function g extending f such that $g(x) = q_0$.

Note that we have implicitly used the fact that the operator $+$ is commutative and associative, up to bisimulation (see the equations in Table 2). Note also that the second rule is actually not needed: We added it just to associate a finite process to an acyclic automaton.

Proposition 4 *Let $\langle \mathcal{A}, f \rangle$ be an automaton over $X_{\mathcal{A}}$ and let Exp$(\langle \mathcal{A}, f \rangle)$ be its essentially unique process. Then $\mathcal{A}$ is bisimilar to LTS(Exp$(\langle \mathcal{A}, f \rangle)$).*

The proof can be given by coinductive arguments, by associating to the root of $\mathcal{A}$ the state *Exp*$(\langle \mathcal{A}, f \rangle)$, and to each state q_i all the processes *Exp*$(\langle \mathcal{A}_i, g \rangle)$ arising during the translation, and such that q_i is the root of $\mathcal{A}_i$.

Example 3 *Consider now the two component automata $\mathcal{C}_1 = (\{p, p'\}, \{b\}, \{(p, b, p')\}, \{p\})$ and $\mathcal{C}_2 = (\{q, q'\}, \{a, b\}, \{(q, b, q), (q, a, q')\}, \{q\})$ from Example 1 as automata over the sets of state variables $X_{\mathcal{C}_1}$ and $X_{\mathcal{C}_2}$, respectively.*

By Definition 9, Exp$(\langle \mathcal{C}_1, f_1 \rangle) = b$.Exp$(\langle \mathcal{C}'_1, f_1 \rangle)$, with $\mathcal{C}'_1 = (\{p'\}, \{b\}, \varnothing, p')$, and Exp$(\langle \mathcal{C}'_1, f_1 \rangle) = \mathsf{nil}$; thus Exp$(\langle \mathcal{C}_1, f_1 \rangle) = b.\mathsf{nil}$. Moreover, $\mathcal{C}_1$ trivially is bisimilar to LTS$(b.\mathsf{nil})$.

Again by Definition 9, Exp$(\langle \mathcal{C}_2, f_2 \rangle) = rec_x.(\, b.x + a$.Exp$(\langle \mathcal{C}'_2, f'_2 \rangle)\,)$, with $\mathcal{C}'_2 = (\{q, q'\}, \{a, b\}, \varnothing, q')$ and $f'_2(x) = q$, and Exp$(\langle \mathcal{C}'_2, f'_2 \rangle) = \mathsf{nil}$; thus Exp$(\langle \mathcal{C}_2, f_2 \rangle) = rec_x.(b.x + a.\mathsf{nil})$. Finally, $\mathcal{C}_2$ trivially is bisimilar to LTS$(rec_x.(b.x + a.\mathsf{nil}))$.

It is worth noting that the encoding presented in Definition 9 can be proved to be compositional: Thus, up to bisimulation, the parallel composition of two automata, according to any of the coordinaton patterns, is mapped into a process that is bisimilar to the parallel composition, according to the corresponding operator, of the encoding of the underlying automata.

Proposition 5 *Let $\mathcal{A}$ and $\mathcal{B}$ be automata, and $\langle \mathcal{A}, \varnothing \rangle$ and $\langle \mathcal{B}, \varnothing \rangle$ the associated automata over an empty set of state variables. Then $\mathcal{A} \parallel_L \mathcal{B}$ is bisimilar to LTS(Exp$(\langle \mathcal{A}, \varnothing \rangle)) \parallel_L$ LTS(Exp$(\langle \mathcal{B}, \varnothing \rangle))$ for any set of names L and any of the three parallel operators.*

5. Equations for (finite) languages

Consider again the equational presentation for bisimulation offered in Section 3. In particular, note how the normal form of a finite process intuitively corresponds to a regular expression, obtained by using the set of actions of the calculus as the alphabet and action prefixing and non-deterministic choice as operations. This intuition can be exploited to obtain an equational presentation for the language of a team automaton.

The correspondence between regular expressions and languages is a staple of theoretical computer science, so we do not repeat it here. We simply let $\mathcal{L}_P$ denote the language of a process P, which is easily derived from its normal form. Moreover, we let $\widehat{\mathcal{L}}$ denote the prefix-closed extension of a language $\mathcal{L}$ over Σ, i.e.,

$$\widehat{\mathcal{L}} = \{\, \alpha \in \Sigma^* \mid \exists\, \beta \in \Sigma^* :\ \alpha\beta \in \mathcal{L} \,\}.$$

As a direct corollary of Proposition 4 we thus obtain the following result.

Proposition 6 *Let $\mathcal{A}$ be an automaton. Then $\mathbb{L}(\mathcal{A}) = \widehat{\mathcal{L}}_{Exp(\langle \mathcal{A}, \varnothing \rangle)}$.*

This result suggests that our calculus can be used to derive the language of an automaton. This is not surprising, since bisimulation is finer than language equivalence, even if the environment is slightly more complex than usual, since our calculus contains three different operators for parallel composition. This result moreover suggests the use of equational laws to distill a normal form that is simpler than the original automaton.

Proposition 7 *Let P, Q be finite processes. Then $\widehat{\mathcal{L}}_P = \widehat{\mathcal{L}}_Q$ if and only if the normal forms of P and Q are equated by using the axioms of $+$ (except for idempotency, see Table 2) and the axiom*

$$a.M + a.N = a.(M + N).$$

Also this equation can be interpreted as a left-to-right rewriting rule, allowing for further reduced normal forms of processes. It is important to realize that this axiom could not simply have been added to Tables 2 and 3, since *critical pairs* would have arisen due to this axiom's incompatibility with the distributivity of eager parallel composition.

Example 4 *Consider the three automata $\mathcal{D}_1$, $\mathcal{D}_2$ and $\mathcal{D}_3$ used for providing the counterexample concerning Proposition 1, as shown in Figure 3. If we ignore the above axiom, then clearly their associated processes D_1, D_2 and D_3 have the normal forms $a.b.\mathsf{nil}$, $a.b.\mathsf{nil} + a.c.\mathsf{nil}$ and $a.(b.\mathsf{nil} + c.\mathsf{nil})$, respectively. Should the above axiom have been added to the set of equations in Tables 2 and 3, then clearly D_2 would be equated to D_3 and thus $D_1 \parallel_\Sigma^{si} D_2$ would have the same normal form (hence recognize the same langauge) as $D_1 \parallel_\Sigma^{si} D_3$, which is not the case: Instead, the normal form for $a.b.\mathsf{nil} \parallel_\Sigma^{si} a.b.\mathsf{nil} + a.c.\mathsf{nil}$ is $a.b.\mathsf{nil} + a.b.c.\mathsf{nil} + a.c.b.\mathsf{nil}$, reduced to $a.(b.c.\mathsf{nil} + c.b.\mathsf{nil})$; while the normal form for $a.b.\mathsf{nil} \parallel_\Sigma^{si} a.(b.\mathsf{nil} + c.\mathsf{nil})$ is $a.b.\mathsf{nil} + a.c.b.\mathsf{nil}$, reduced to $a.(b.\mathsf{nil} + c.b.\mathsf{nil})$. The associated languages are easily derived.*

The situation so far is thus quite satisfactory for finite processes (i.e., equivalently, for acyclic automata): In order to prove the equivalence of two team automata with respect to the language they recognize, it is sufficient to consider the associated processes and analyze their normal forms. Moreover, it is relevant that the mapping from team automata to processes preserves, up to bisimulation, the three composition patterns that were consideres in this paper: This result ensures that the procedure devised so far for obtaining the normal form is modular.

6. Conclusions and future work

We introduced a process calculus for team automata, extending some classical results on I/O automata. As a side-effect, we widened the family of team automata that guarantee a degree of compositionality by providing a way to obtain the language of a (finite) $\mathcal{R}^{si}$-team automaton from its components. While this language cannot be obtained through a direct manipulation of the languages of the component automata, the resulting degree of modularity favors the use of team automata in component-based system design.

Future work in this direction should lead to compositionality results for other types of team automata. A first step in this direction could be to extend our calculus with parallel composition operators that mimic the various peer-to-peer and master-slave patterns of synchronization for team automata as introduced in [ter Beek et al. 2003], as well as mixtures of the synchronizations defined for team automata. As a matter of fact, [ter Beek 2003, ter Beek and Kleijn 2003] contain compositionality results not only for $\mathcal{R}^{free}$- and $\mathcal{R}^{ai}$-team automata, but also for team automata constructed according to a mixture of the *free* and *ai* synchronizations. It is important to recall, however, that the various peer-to-peer and master-slave patterns of synchronization make use of the distinction of the set of actions of team automata into input, output and internal actions. This means that in order to tackle the above issues, our calculus should first be extended to take this distinction into account.

Our correspondence results between automata and processes (as summed up by the two propositions in Section 4) relate the behavior of *possibly cyclic* automata and *possibly recursive* processes. We restrained however from tackling the axiomatization of recursive processes. It would be relatively easy to come up with a complete set of equational laws for those recursive

processes not containing the parallel operators, since they basically boil down to regular expressions equipped with a Kleene star operator. On the other hand, the lack of a complete set of axioms for recursive processes is a common trait for all the calculi proposed for automata that we are aware of (compare, e.g., the situation for (probabilistic) I/O automata, as reported in [De Nicola and Segala 1995, Stark et al. 2003]). We hope that our syntactical restriction will suffice to obtain a relatively small set of equations which is complete, but we leave this topic as the subject of future work.

Lastly, in order to be really useful in practical applications of team automata, it would be worthwhile to study the complexity of the algorithms introduced in this paper, e.g., what is the cost of obtaining the language of a team automaton via its translation into a process.

Acknowledgments

We are grateful to Jetty Kleijn and to the anonymous referees for their useful comments on a preliminary version of this paper.

References

Aceto, L., Bloom, B., and Vaandrager, F. W. (1994). Turning SOS rules into equations. *Information and Computation*, 111(1):1–52.

ter Beek, M. H. (2003). *Team Automata—A Formal Approach to the Modeling of Collaboration Between System Components*. PhD thesis, Leiden University.

ter Beek, M. H., Ellis, C. A., Kleijn, J., and Rozenberg, G. (2003). Synchronizations in team automata for groupware systems. *Computer Supported Cooperative Work*, 12(1):21–69.

ter Beek, M. H. and Kleijn, J. (2003). Team automata satisfying compositionality. In Araki, K., Gnesi, S., and Mandrioli, D., editors, *International Symposium of Formal Methods Europe*, volume 2805 of *Lect. Notes in Comp. Sci.*, pages 381–400. Springer.

ter Beek, M. H. and Kleijn, J. (2005). Modularity for teams of I/O automata. *Information Processing Letters*, 95(5):487–495.

ter Beek, M. H., Lenzini, G., and Petrocchi, M. (2005). Team automata for security: A survey. In R., F. and Zavattaro, G., editors, *International Workshop on Security Issues in Coordination Models, Languages, and Systems*, volume 128 of *Electr. Notes in Theor. Comp. Sci.*, pages 105–119. Elsevier.

ter Beek, M. H., Lenzini, G., and Petrocchi, M. (2006). A team automaton scenario for the analysis of security properties of communication protocols. *Journal of Automata, Languages and Combinatorics*. To appear.

Bergstra, J. A. and Klop, J. W. (1984). Process Algebra for Synchronous Communication. *Information and Control*, 60(1–3):109–137.

Bol, R. N. and Groote, J. F. (1996). The meaning of negative premises in transition system specifications. *Journal of the ACM*, 43(5):863–914.

De Nicola, R. and Segala, R. (1995). A process algebraic view of input/output automata. *Theoretical Computer Science*, 138(2):391–423.

Ellis, C. A. (1997). Team automata for groupware systems. In *International Conference on Supporting Group Work*, pages 415–424. ACM Press.

Hoare, C. A. R. (1985). *Communicating Sequential Processes*. Prentice Hall.

Jonsson, B. (1994). Compositional specification and verification of distributed systems. *ACM Transactions on Programming Languages and Systems*, 16(2):259–303.

Kleijn, J. (2003). Team automata for CSCW: A survey. In *Petri Net Technology for Communication-Based Systems*, volume 2472 of *Lect. Notes in Comp. Sci.*, pages 295–320. Springer.

Lynch, N. A. (1996). *Distributed Algorithms*. Morgan Kaufmann.

Lynch, N. A. and Tuttle, M. R. (1989). An introduction to input/output automata. *CWI Quarterly*, 2(3):219–246. Also appeared as Technical Memo MIT/LCS/TM-373, MIT, 1988.

Milner, R. (1980). *A Calculus of Communicating Systems*, volume 92 of *Lect. Notes in Comp. Sci.* Springer.

Plotkin, G. (1981). A structural approach to operational semantics. Technical Report DAIMI FN-19, Computer Science Department, Aarhus University.

Stark, E. W., Cleaveland, R., and Smolka, S. A. (2003). A process-algebraic language for probabilistic I/O automata. In Amadio, R. and Lugiez, D., editors, *International Conference on Concurrency Theory*, volume 2761 of *Lect. Notes in Comp. Sci.*, pages 193–207. Springer.

Vaandrager, F. W. (1991). On the relationship between process algebra and input/output automata (extended abstract). In *Symposium on Logic in Computer Science*, pages 387–398. IEEE Computer Society Press.