

Investigating Fluid-Flow Semantics of Asynchronous Tuple-Based Process Languages for Collective Adaptive Systems

Diego Latella¹, Michele Loreti², and Mieke Massink¹(✉)

¹ Istituto di Scienza e Tecnologie dell'Informazione 'A. Faedo', CNR, Pisa, Italy
{Diego.Latella,mieke.massink}@isti.cnr.it

² Università di Firenze and IMT-Lucca, Lucca, Italy
michele.loreti@unifi.it

Abstract. Recently, there has been growing interest in nature-inspired interaction paradigms for Collective Adaptive Systems, for modelling and implementation of adaptive and context-aware coordination, among which the promising pheromone-based interaction paradigm. System modelling in the context of such a paradigm may be facilitated by the use of languages in which adaptive interaction is decoupled in time and space through asynchronous buffered communication, e.g. asynchronous, repository- or tuple-based languages. In this paper we propose a differential semantics for such languages. In particular, we consider an asynchronous, repository based modelling kernel-language which is a restricted version of LINDA, extended with stochastic information about action duration. We provide stochastic formal semantics for both an agent-based view and a population-based view. We then derive an ordinary differential equation semantics from the latter, which provides a fluid-flow deterministic approximation for the mean behaviour of large populations. We show the application of the language and the ODE analysis on a benchmark example of foraging ants.

Keywords: Asynchronous coordination languages · Stochastic process algebras · Fluid-flow approximation · Continuous time markov chains

1 Introduction and Related Work

Collective Adaptive Systems (CAS) are systems typically composed of a large number of heterogeneous agents with decentralised control and varying degrees of complex autonomous behaviour. Agents may be competing for shared resources and, at the same time, collaborate for reaching common goals. The pervasive nature of CAS, together with the importance of the role they play, for instance in the very core of the ICT support for *smart cities*, implies that a serious *a priori* analysis—and, consequently modelling—of the design of any such a system must be performed and that all critical aspects of its behaviour must be carefully investigated before the system is deployed.

This research has been partially funded by the EU project QUANTICOL (nr. 600708), and the IT MIUR project CINA.

Recently, there has been growing interest in nature-inspired interaction paradigms for CAS, for enforcing adaptive and context-aware coordination. Among these, those based on the metaphor of pheromones seem promising. System modelling in the context of such a paradigm may be facilitated by the use of languages in which adaptive interaction is decoupled in time and space through asynchronous buffered communication, e.g. tuple-based languages, *a la* LINDA [5]. For systems of limited size, several languages have already been proposed in the literature and have proven useful for modelling—as well as programming—autonomic adaptive coordination. Examples include KLAIM [10], which extends LINDA with, among others, a notion of *space*, the TOTA framework [21], which, additionally, provides for explicit adaptive *tuple propagation* mechanisms and a sort of force field view of tuples, and SCEL [8], where the basic interaction paradigm is enriched with a flexible, *predicate-based* addressing mechanism, with a framework for defining *policies*, and with a notion of tuple-space which is extended to a more general *knowledge-space*. Additionally, quantitative extensions of both KLAIM and SCEL have been developed, namely StoKLAIM [11] and StocS [20], where the quantity of interest is the duration of (the execution of) process actions. Such durations are assumed to be continuous random variables with negative exponential distributions, commonly used in stochastic process algebra [17]. Consequently, each such random variable is fully characterised by its rate, a positive real value that is equal to the inverse of the mean duration of the execution of the action. This choice for action durations gives rise to a Markovian semantics for the languages: the behaviour of each agent of a system is modelled by a continuous time Markov chain (CTMC). The collective behaviour of a system of agents is also modelled by a CTMC, of course obtained as a suitable combination of those of the component agents.

Unfortunately, as soon as the size of the systems under consideration grows, the infamous combinatorial state space explosion problem makes system modelling and analysis essentially unfeasible. On the other hand, one of the key features of CAS is the large size of their component populations. Consequently, *scalability* of modelling—and, most of all, analysis—techniques and tools becomes a *must* in the context of CAS design and development. It is thus essential to develop alternative approaches for modelling systems with *large populations of agents*, possibly based on—and formally linked to—process algebra. In this way, one can try to extend, to such alternative approaches, modelling and analysis techniques which have proven effective for standard stochastic process algebra, such as stochastic model-checking of probabilistic temporal logics. One way to deal with large population systems is the so called *fluid-flow* approach, which consists in computing a deterministic approximation of the mean behaviour of the large population [2]. The first step is to abstract from agent identity and to look only at the *number of agents* in a particular state, for each of the possible states of the agents in the population and at any point in time. Then, a further step is performed by approximating the average values of such numbers by means of a *deterministic, continuous* function of time, which is obtained as the solution of an initial value problem where the set of ordinary differential equations (ODE)

is derived from the system model and the initial condition is the initial distribution of the population over the set of local states of the agents. Prominent examples of the fluid-flow approach are the *differential* (i.e. ODE-based) semantics version of the Performance Analysis Process Algebra (PEPA) [23], which we will call ODE-PEPA, Bio-PEPA [7] and, more recently, PALOMA [13]. The advantage of a *fluid-flow* approach is that the transient average behaviour of the system can be analysed orders of magnitude faster than by stochastic simulation, where the mean of a usually large number of simulation traces must be computed. The *fluid-flow* approach is independent of the size of the involved populations, as long as this size is large enough to provide a good deterministic approximation [2].

In this paper we explore the possibility for differential semantics for languages with an asynchronous buffered communication interaction paradigm, e.g. data-repository- /tuple- based ones. We present ODELINDA, a simple experimental language, based on a LINDA-like, asynchronous paradigm, where processes interact only via a data repository by means of **out**, **in** and **read** operations for respectively inserting, withdrawing and reading data values to/from the repository. In particular, we present a *quantitative*, Markovian language; the behaviour of each agent is modelled by a Markov process.

In most stochastic process languages, each action is decorated with its rate, which is typically a constant. In ODELINDA, instead, action rates are allowed to depend on the global state of the complete system; thus they are *functions* from global system states to positive real values, in a similar way as in Bio-PEPA. We provide a formal definition of the Markovian semantics using State-to-Function Labeled Transition Systems (FuTS) [9], an approach that provides for a simple and concise treatment of transition multiplicities—an issue closely related to the CTMC principle of race-condition—and a clean and compact definition of the semantics.

We follow the fluid-flow approach for making the language scalable in order to be able to deal with CAS. We define a population semantics for ODELINDA from which a differential (ODE) semantics is derived, in a similar way as proposed in [13] for PALOMA and in [23] for ODE-PEPA. The interaction paradigm underlying ODELINDA is fundamentally different from those of ODE-PEPA, Bio-PEPA and PALOMA. ODE-PEPA is based on the well-known PEPA process interaction paradigm, with processes synchronising on common, specific activities, Bio-PEPA is based on the chemical-reaction paradigm, whereas PALOMA agents use message multicasting. Additionally, both Bio-PEPA and PALOMA provide some simple means for spatial modelling. Spatial information is currently not incorporated in ODELINDA.

In tuple-based approaches, data repositories are typically *multi-sets* of values and adding/withdrawing a value to/from the repository increases/decreases the multiplicity of that value in the repository. In a “population”-oriented view, this means that the total system population size may change during the computations, i.e. we are dealing with a birth-death type of systems. This is the case for ODELINDA and constitutes another distinguishing feature when compared with

e.g. ODE-PEPA. In this respect, our proposal is more similar to sCCP [3], although, from a technical point of view, for the actual definition of the differential semantics we followed the approach used in [23,13] rather than that presented in [3]. Finally, our work is also related to PALPS [1] and MASSPA [15]. PALPS is a language for ecological models. Only an individual-based semantics is available for PALPS. The language is thus usable only in the specific domain of ecological models and, furthermore, seriously suffers of lack of scalability. MASSPA [15] shares some features with PALOMA, e.g. a multicast-like interaction paradigm; it is lacking a Markovian, individual-based semantics.

It is worth noting that the language we present here is a minimal *kernel* language; we intended to address only the basic issues which arise when defining a differential semantics for tuple-based asynchronous languages. For this reason, operations on data, and in particular *templates* and *pattern-matching* are not considered, so that **in** and **read** operations result into pure synchronisation actions (with or without value consumption, respectively). The unconstrained use of templates and pattern-matching, as well as the use of general operations on data types, could result in an unbounded number of *distinct* values in a model, which, in turn, would require an unbounded number of differential equations in the differential semantics. Consequently, only ground terms are allowed in model specifications. It is worth noting that this does *not* imply that we allow only finite computations or that there are bounds on the multiplicity of each piece of data or on the resulting state spaces. In fact, the number of copies of any given value which can be stored in a repository by means of repeated executions of **out** actions in a computation, by one or more processes, is unbounded (and may be infinite for infinite computations). Anyway, one should also keep in mind that ODELINDA is intended to be a process *modelling*, rather than a programming language and that differently from most process modelling languages, that, typically, do not provide any feature for dealing with data, offers some means, although primitive, for data storage, withdrawal and retrieval. For the sake of simplicity, we also refrain from considering process spawning, although this would not cause particular problems given that the semantic model we use deals with dynamic population sizes in a natural way. The objective of the present paper is to show that the basic notion of ODE semantics for asynchronous, shared-repository based languages is well founded. Additionally, by revisiting the benchmark example of Foraging Ants, we show that even in the restricted form we present in this paper, ODELINDA can be useful for actual system modelling and analysis.

The present paper is organised as follows: the syntax and Markovian, individual-based semantics of ODELINDA are presented in Section 2; the differential semantics of the language are presented in Section 3. An example of model specification as well as ODE analysis is given in Section 4. Finally, some conclusions and considerations for future work are discussed in Section 5.

2 Syntax and Markovian Semantics of OdeLinda

We recall that the main purpose of this paper is to show the basic principles for the definition of a differential semantics of asynchronous repository-based languages rather than the definition of a complete, high-level, ready-to-use process language. Consequently, the language we present here is a very minimal one, although, as we pointed out in Section 1 it can be used for the effective modelling of typical CAS systems like foraging ants, as we will show in Section 4.

2.1 Syntax

Let \mathcal{D} be a denumerable non-empty set of *data* values, ranged over by d, d', d_1, \dots , \mathcal{A} be set of *actions* with $\mathcal{A} = \mathcal{A}_o \cup \mathcal{A}_i \cup \mathcal{A}_r$, where $\mathcal{A}_o = \{\mathbf{out}(d) \mid d \in \mathcal{D}\}$, $\mathcal{A}_i = \{\mathbf{in}(d) \mid d \in \mathcal{D}\}$, $\mathcal{A}_r = \{\mathbf{read}(d) \mid d \in \mathcal{D}\}$, ranged over by α, α', \dots , \mathcal{P} be a denumerable non-empty set of *state constants* (or states), ranged over by C, C', C_1, \dots .

A system model is the result of the *parallel composition of agents*, i.e. *processes*, which are finite state machines. Thus the language has the following two level grammar for the sets AGENTS of agents and PROC of processes:

$$A ::= (R, \mathbf{out}(d)).C \mid (R, \mathbf{in}(d)).C \mid (R, \mathbf{read}(d)).C \mid A + A \quad P ::= C \mid P \parallel P$$

where for each used constant C there is a definition of the form $C := A$, which, in the sequel, will be written as $C := \sum_{j \in J} (R_j, \alpha_j).C_j$, for some finite index set J , with obvious meaning. In action prefix, $(R, \alpha)_{\dots}$, R is the name of a *rate function* under the scope of a suitable definition $R := E$; E is a numerical expression where the special operator $\#C$ can be used which, for state name C , yields the number of agents which are in state C in the current global system state. We will refrain from giving further details on the syntax of expressions E .

A *process definition* is the collection of definitions for the states of the process. A *system state* is a pair (P, D) where the set REPS of data repositories D is defined according to the following grammar:

$$D ::= \langle \rangle \quad \mid \quad \langle d \rangle \quad \mid \quad D \mid D$$

The language of expressions E for rate function definitions is extended with $\#d$, for values $d \in \mathcal{D}$, with the obvious meaning. A *system (model) specification* is composed of the set of definitions for its processes, the set of definitions for the rate functions used therein, and an initial global state (P_0, D_0) . It is required that for each state in the system specification there is exactly one definition. For the sake of simplicity, in the present paper we require that for all $i \in I$ and $x \in \{o, i, r\}$, if $\alpha_{ij} \in \mathcal{A}_x$ for some $j \in J_i$, then $\alpha_{ih} \in \mathcal{A}_x$ for all $h \in J_i$ (no mixed choice) and that for $C \neq C'$, if $C := \sum_{j \in J} (R_j, \alpha_j).C_j$ and $C' := \sum_{j \in J'} (R'_j, \alpha'_j).C'_j$ are both state definitions appearing in the system definition then $\{\alpha_j\}_{j \in J} \cap \{\alpha'_j\}_{j \in J'} \neq \emptyset$ implies $\{\alpha_j\}_{j \in J} = \{\alpha'_j\}_{j \in J'}$ (in order not to incur in the possibility of circular definitions in the ODE). In the sequel we let \mathcal{S} denote the set of global system states.

Agents:	
Reader	= (RA, in (a)).Comp + (RB, in (b)).Comp
Comp	= (RR, read (r)).Reader
AWriter	= (WA, out (a)).AWriter
BWriter	= (WB, out (b)).AWriter
Rate Functions:	
RA	= $10 \cdot \#Reader \cdot \#a$
RB	= $5 \cdot \#Reader \cdot \#b$
RR	= $10 \cdot \#Comp \cdot \#r$
WA	= $9 \cdot \#AWriter$
WB	= $4 \cdot \#BWriter$

Fig. 1. A simple model of Readers and Writers

As a simple running example we consider the specification of a readers/writers model given in Figure 1, where two kinds of writers are considered—those writing messages of type a and those writing messages of type b —and readers perform some computation using some resources r before reading the next item, modelled by synchronisation on r —with the following initial state¹

$(Reader[10000] \parallel AWriter[5000] \parallel BWriter[5000], \langle a \rangle[5000] | \langle b \rangle[5000]) \langle r \rangle[1000]$.

2.2 Stochastic Semantics

The stochastic semantics are given in Figure 2 using the FuTS framework [9], that is an alternative to the classical approach, based on Labelled Transition Systems (LTS). In LTS, a transition is a triple (s, α, s') where s and α are the source state and the label of the transition, respectively, while s' is the target state reached from s via a transition labeled with α . In FuTS, a transition is a triple of the form (s, α, \mathcal{F}) . The first and second component are the source state and the label of the transition, as in LTS, while the third component \mathcal{F} is a *continuation function* (or simply a *continuation* in the sequel), which associates a value from an appropriate semiring with each state s' . In the case of Markovian process algebra, the relevant semiring is that of non-negative real numbers. If \mathcal{F} maps s' to 0, then state s' cannot be reached from s via this transition. A positive value for state s' represents the rate for the jump of the system from s to s' . Any FuTS over $\mathbb{R}_{\geq 0}$ uniquely defines a CTMC, which can obviously be built by successive application of the continuations to the set of states. Below we recall the main notions on FuTS. The reader interested in further details is referred to [9].

Given a denumerable non-empty set V , we let $\mathbf{FS}(V, \mathbb{R}_{\geq 0})$ denote the class of finitely supported² functions from V to $\mathbb{R}_{\geq 0}$. For v_1, \dots, v_n in set V and

¹ We use the standard notational convention that $P[n]$ means n instances of P in parallel: $P \parallel P \parallel \dots \parallel P$. We extend it to tuples in the obvious way.

² A function $f : V \rightarrow \mathbb{R}_{\geq 0}$ has finite support if and only if the set $\{v \in V \mid f v \neq 0\}$ is finite. In this paper we often use Currying in function application.

$$\begin{array}{c}
\text{PA: } \frac{C := \sum_{j \in J} (R_j, \alpha_j) \cdot C_j}{C \xrightarrow{R, \alpha} \sum_{\{h \in J, \alpha_h = \alpha \wedge R_h = R\}} [C_h \mapsto 1]} \quad \text{DI: } \frac{}{\langle d \rangle \xrightarrow{\text{in}(d)} [\langle \rangle \mapsto 1]} \\
\\
\text{DR: } \frac{}{\langle d \rangle \xrightarrow{\text{read}(d)} [\langle d \rangle \mapsto 1]} \quad \text{DN}_1: \frac{}{\langle \rangle \xrightarrow{\alpha} []} \quad \text{DN}_2: \frac{\alpha \notin \{\text{in}(d), \text{read}(d)\}}{\langle d \rangle \xrightarrow{\alpha} []} \\
\\
\text{PP: } \frac{\frac{P_1 \xrightarrow{R, \alpha} \mathcal{P}_1 \quad P_2 \xrightarrow{R, \alpha} \mathcal{P}_2}{P_1 | P_2 \xrightarrow{R, \alpha} \mathcal{P}_1 | (\mathcal{X} P_2) + (\mathcal{X} P_1) | \mathcal{P}_2}}{} \quad \text{DP: } \frac{D_1 \xrightarrow{\alpha} \mathcal{P}_1 \quad D_2 \xrightarrow{\alpha} \mathcal{P}_2}{D_1 | D_2 \xrightarrow{\alpha} \mathcal{P}_1 | (\mathcal{X} D_2) + (\mathcal{X} D_1) | \mathcal{P}_2} \\
\\
\text{OUT: } \frac{P \xrightarrow{R, \text{out}(d)} \mathcal{P}}{(P, D) \xrightarrow{\text{out}(d)} \mathcal{R}_o(P, D, \mathcal{P}, \mathcal{X}(D | \langle d \rangle), R)} \\
\\
\text{IN: } \frac{D \xrightarrow{\text{in}(d)} \mathcal{P} \quad P \xrightarrow{R, \text{in}(d)} \mathcal{P}}{(P, D) \xrightarrow{\text{in}(d)} \mathcal{R}_i(P, D, \mathcal{P}, \mathcal{P}, R)} \quad \text{READ: } \frac{D \xrightarrow{\text{read}(d)} \mathcal{P} \quad P \xrightarrow{R, \text{read}(d)} \mathcal{P}}{(P, D) \xrightarrow{\text{read}(d)} \mathcal{R}_r(P, D, \mathcal{P}, \mathcal{P}, R)}
\end{array}$$

Fig. 2. FuTS semantics of the process language with tuple creation

$r_1, \dots, r_n \in \mathbb{R}_{\geq 0}$, we let $[v_1 \mapsto r_1, \dots, v_n \mapsto r_n]$ in $\mathbf{FS}(V, \mathbb{R}_{\geq 0})$ denote the function mapping v_i to r_i , for $i = 1, \dots, n$, and any other $v \in V \setminus \{v_1, \dots, v_n\}$ to 0; the degenerate case $[]$ denotes the constant function yielding 0 everywhere. For $v \in V$, $\mathcal{X}v$ denotes the function $[v \mapsto 1]$. For functions $\mathcal{F}_1, \mathcal{F}_2 \in \mathbf{FS}(V, \mathbb{R}_{\geq 0})$ we let $(\mathcal{F}_1 + \mathcal{F}_2) \in \mathbf{FS}(V, \mathbb{R}_{\geq 0})$ be defined as $(\mathcal{F}_1 + \mathcal{F}_2)v = (\mathcal{F}_1 v) + (\mathcal{F}_2 v)$ and we extend $(\mathcal{F}_1 + \mathcal{F}_2)$ to the n -ary version $\sum_{j \in J} \mathcal{F}_j$, in the obvious way, for finite index set J . For $r \in \mathbb{R}$ we let $\mathcal{F}/r \in \mathbf{FS}(V, \mathbb{R}_{\geq 0})$ be the defined as $(\mathcal{F}/r)v = (\mathcal{F}v)/r$ if $r \neq 0$ and $(\mathcal{F}/r)v = 0$ otherwise. We let $\oplus \mathcal{F}$ be defined as $\oplus \mathcal{F} = \sum_{v \in V} (\mathcal{F}v)$; note that $\oplus \mathcal{F}$ is finite, and thus well-defined, for $\mathcal{F} \in \mathbf{FS}(V, \mathbb{R}_{\geq 0})$. We recall standard structural congruence \equiv on REPS, with $D | \langle \rangle \equiv D, D_1 | D_2 \equiv D_2 | D_1, (D_1 | D_2) | D_3 \equiv D_1 | (D_2 | D_3)$. In the sequel, when dealing with data repositories, we will implicitly assume them modulo \equiv . For the sake of notational simplicity we will keep $D, D' \dots$ in the notation (but actually the representatives of their equivalence classes are intended). A similar structural congruence \equiv is assumed for processes, with $P_1 \parallel P_2 \equiv P_2 \parallel P_1, (P_1 \parallel P_2) \parallel P_3 \equiv P_1 \parallel (P_2 \parallel P_3)$, as well as similar conventions concerning notation.

For function \mathcal{P} in $\mathbf{FS}(\text{PROC}_{/\equiv}, \mathbb{R}_{\geq 0})$ and \mathcal{D} in $\mathbf{FS}(\text{REPS}_{/\equiv}, \mathbb{R}_{\geq 0})$ the notation $(\mathcal{P}, \mathcal{D})$ defines a function in $\mathbf{FS}(\mathcal{S}_{/\equiv}, \mathbb{R}_{\geq 0})$ as follows: for system state $(P, D) \in \mathcal{S}_{/\equiv}$, we have $(\mathcal{P}, \mathcal{D})(P, D) = (\mathcal{P}P) \cdot (\mathcal{D}D)$, where \cdot denotes product in $\mathbb{R}_{\geq 0}$. For each rate function definition $R = E$, we consider the function $R : \mathcal{S}_{/\equiv} \rightarrow \mathbb{R}_{\geq 0}$ defined in the following. For all $(P, D) \in \mathcal{S}_{/\equiv}$ $R(P, D) = \llbracket E \rrbracket_{(P, D)}$, where $\llbracket E \rrbracket_{(P, D)}$ denotes the value of expression E in the current global state (P, D) . Obviously $\llbracket \#C \rrbracket_{(C, D)} = 1$ and $\llbracket \#C \rrbracket_{(P_1 \parallel P_2, D)} = \llbracket \#C \rrbracket_{(P_1, D)} + \llbracket \#C \rrbracket_{(P_2, D)}$. The definition for $\llbracket \#d \rrbracket_{(P, D)}$ is similar.

The continuation summation in Rule (PA) takes care of multiple alternatives with the same action and rate function. Different choices for functions

\mathcal{R}_o , \mathcal{R}_i , \mathcal{R}_r in rules (IN), (OUT) and (READ) give rise to different interaction policies. For instance, for a TIPP-like synchronisation policy, assuming each rate function definition be of the form $R := k_R$, for $k_R \in \mathbb{R}_{>0}$, one can let $\mathcal{R}_o(P, D, \mathcal{P}, \mathcal{D}, R) = \mathcal{R}_i(P, D, \mathcal{P}, \mathcal{D}, R) = \mathcal{R}_r(P, D, \mathcal{P}, \mathcal{D}, R) = R(P, D) \cdot (\mathcal{P}, \mathcal{D})$. Similarly, for a PEPA-like interaction paradigm, assuming again each rate function be a constant, we get $\mathcal{R}_o(P, D, \mathcal{P}, \mathcal{D}, R) = R(P, D) \cdot (\mathcal{P}, \mathcal{D})$, and $\mathcal{R}_i(P, D, \mathcal{P}, \mathcal{D}, R) = \mathcal{R}_r(P, D, \mathcal{P}, \mathcal{D}, R) = R(P, D) \cdot \frac{\min\{\oplus\mathcal{P}, \oplus\mathcal{D}\}}{\oplus\mathcal{P} \cdot \oplus\mathcal{D}}(\mathcal{P}, \mathcal{D})$. In this paper we choose $\mathcal{R}_o = \mathcal{R}_i = \mathcal{R}_r = \mathcal{R}$ where

$$\mathcal{R}(P, D, \mathcal{P}, \mathcal{D}, R) = R(P, D) \cdot \left(\frac{\mathcal{P}}{\oplus\mathcal{P}}, \frac{\mathcal{D}}{\oplus\mathcal{D}} \right).$$

The idea is that the rate of the action is the full responsibility of the modeller, being equal to $R(P, D)$; in fact $\left(\frac{\mathcal{P}}{\oplus\mathcal{P}}, \frac{\mathcal{D}}{\oplus\mathcal{D}} \right)(P', D')$ is equal to 1 if (P', D') is reachable in one transition from (P, D) and 0, if it is not.

3 Differential Semantics of OdeLinda

In this section we define the differential semantics for the language introduced in Sect. 2. We follow a similar approach as in [13,23]: we first define a population semantics for the language and then we define the differential semantics by means of deriving, from the population semantics, suitable ODEs for the mean-field model.

3.1 Population Semantics

Assume we are given a system specification where $\{C_1, \dots, C_s\}$ is the set of all states of all processes and $\{d_1, \dots, d_t\}$ is the set of all data values textually occurring in the specification. Given a global system state (P, D) we consider the corresponding vector $\mathbf{X} = (x_1, \dots, x_m)$ of counting variables such that, for $i = 1, \dots, s$, x_i records the number of agents in P which are in (local) state C_i , and for $i = s + 1, \dots, m = s + t$, x_i records the number of instances of d_{i-s} in D . Clearly, every transition at the single agent level corresponds to a change in the value of \mathbf{X} , i.e. a *population-based* transition. In order to formalise how single agent transitions induce population-based transitions, let (P', D') and (P'', D'') be two global system states with $P' = C'_1 \parallel \dots \parallel C'_{s'}$, $P'' = C''_1 \parallel \dots \parallel C''_{s''}$, $D' = d'_1 \dots |d'_{t'}$, and $D'' = d''_1 \dots |d''_{t''}$ and, with reference to the given system specification, define the *update vector* δ in the usual way³: $\delta((P'', D''), (P', D')) = (\delta_1, \dots, \delta_m)$ with

$$\delta_i = \begin{cases} \sum_{j=1}^{s''} \mathbf{1}\{C''_j = C_i\} - \sum_{j=1}^{s'} \mathbf{1}\{C'_j = C_i\}, & \text{for } i = 1, \dots, s \\ \sum_{j=1}^{t''} \mathbf{1}\{d''_j = d_i\} - \sum_{j=1}^{t'} \mathbf{1}\{d'_j = d_i\}, & \text{for } i = s + 1, \dots, m \end{cases}$$

³ $\mathbf{1}\{C = C'\}$ is equal to 1 if $C = C'$ and to 0 otherwise.

With the definition of the update vector in place we can easily define the population-based transitions using the following rule:

$$\frac{(P, D) \xrightarrow{\alpha} (\mathcal{P}, \mathcal{D}) \quad r(P, D) = (\mathcal{P}, \mathcal{D})(P', D') > 0}{\mathbf{X} \xrightarrow{\alpha, r(P, D)} \mathbf{X} + \delta((P', D'), (P, D))}$$

Using the above procedure, for any system specification we can derive a population-based CTMC (PCTMC) [2]. Such a PCTMC is defined as the tuple $(\mathbf{X}, \mathbb{Z}^m, \mathcal{T}, \mathbf{x}_0)$, where, :

- $\mathbf{X} = (x_1, \dots, x_m)$ is the state vector, where, for $i = 1, \dots, s$ element x_i is the count of agents in state C_i and, for $i = s + 1, \dots, m = s + t$ it counts the number of instances of d_{i-s} ;
- $\mathcal{T}(\mathbf{X}) = \{\tau_1, \dots, \tau_h\}$ is the set of population-based transitions enabled in state \mathbf{X} . Each transition τ is associated with a update vector δ_τ and a rate $r_\tau(\mathbf{X}) = \sum \{r \mid \mathbf{X} \xrightarrow{\alpha, r} \mathbf{X} + \delta_\tau \text{ for some } \alpha\}$;
- $\mathbf{x}_0 \in \mathbb{Z}^m$ is the initial state of the PCTMC.

3.2 Mean-Field Model

The dynamics of the above PCTMC is as follows: if the PCTMC is currently in state \mathbf{X} , then, every $1/r_\tau(\mathbf{X})$ time units, on average, a change in the population level of some agents and data items δ_τ occurs. We can approximate such a discrete change in a continuous way so that for small finite time interval Δt the change in the population level is

$$\mathbf{X}(t + \Delta t) = \mathbf{X}(t) + r_\tau(\mathbf{X}(t)) \cdot \Delta t \cdot \delta_\tau$$

from which, for $\Delta t \rightarrow 0$, we get the ODE $\frac{d\mathbf{X}(t)}{dt} = r_\tau(\mathbf{X}(t)) \cdot \delta_\tau$. Taking all enabled transitions into account the ODE describing the approximated transient evolution of the complete population-level system dynamics is given by the initial value problem:

$$\frac{d\mathbf{X}(t)}{dt} = \sum_{k=1}^h r_{\tau_k}(\mathbf{X}(t)) \cdot \delta_{\tau_k} \quad \text{with } \mathbf{X}(0) = \mathbf{x}_0$$

for large populations and under suitable scalability assumptions (on the rate functions); the interested reader is referred to [2] for the technical details.

With reference to our running example of Figure 1 we get the equations of Figure 3. Note that there is no dynamics for the writer processes in this example, since each of these agents has just a single state (and a self-loop). Similarly for the resource r .

$$\begin{aligned}
\frac{d \text{Reader}(t)}{dt} &= 10 \cdot r(t) \cdot \text{Comp}(t) - (10 \cdot a(t) + 5 \cdot b(t)) \cdot \text{Reader}(t) \\
\frac{d \text{Comp}(t)}{dt} &= (10 \cdot a(t) + 5 \cdot b(t)) \cdot \text{Reader}(t) - 10 \cdot r(t) \cdot \text{Comp}(t) \\
\frac{d a(t)}{dt} &= 9 \cdot \text{AWriter}(t) - 10 \cdot a(t) \cdot \text{Reader}(t) \\
\frac{d b(t)}{dt} &= 4 \cdot \text{BWriter}(t) - 5 \cdot a(t) \cdot \text{Reader}(t)
\end{aligned}$$

Fig. 3. ODE for the simple model of Readers and Writers of Figure 1

4 Example - Foraging Ants

As an example, we revisit a somewhat simplified model of a colony of foraging ants inspired by earlier work in the literature [14,12,22]. The ants initially reside at a *Nest* and will move between the *Nest* and a *Food* site. There are two, bidirectional, paths connecting the *Nest* to the *Food* site (and vice-versa), the *Fast* path and the *Slow* path. Each path is composed by a finite sequence of (path) stages: the number ℓ_F of stages of the *Fast* path is smaller than the number ℓ_S of stages of the *Slow* path. The average time it takes an ant to traverse a stage is the same for each stage, regardless of whether it is situated on the *Slow* or the *Fast* path; such traversal times are modelled by exponentially distributed random variables. The situation is depicted in Fig. 4 where FP_j stands for the j -th of the ℓ_F stages of the *Fast* path and SP_j stands for the j -th of the ℓ_S stages of the *Slow* path, for $\ell_F < \ell_S$. A model for foraging ants is specified in Fig. 5. The set of data values occurring in the model specification is the finite set $(\bigcup_{j=1}^{\ell_F} \{\text{Phe@FP}_j\}) \cup (\bigcup_{j=1}^{\ell_S} \{\text{Phe@SP}_j\})$ where tuple $\langle \text{Phe@FP}_j \rangle$ ($\langle \text{Phe@SP}_j \rangle$, respectively) represents a unit of pheromone in stage j of the *Fast* path (*Slow* path, respectively). There are two process types, one modelling an ant and one modelling the expiration, i.e. decay, of pheromones; the (finite) set of states is as follows:

$$\begin{aligned}
&\{\text{Ant@Nest}, \text{Ant@Food}\} \cup \\
&(\bigcup_{j=1}^{\ell_F} \{\text{AntToFood@FP}_j\}) \cup (\bigcup_{j=1}^{\ell_S} \{\text{AntToFood@SP}_j\}) \cup \\
&(\bigcup_{j=1}^{\ell_F} \{\text{AntToNest@FP}_j\}) \cup (\bigcup_{j=1}^{\ell_S} \{\text{AntToNest@SP}_j\}) \cup \\
&(\bigcup_{j=1}^{\ell_F} \{\text{ExpPhe@FP}_j\}) \cup (\bigcup_{j=1}^{\ell_S} \{\text{ExpPhe@SP}_j\}).
\end{aligned}$$

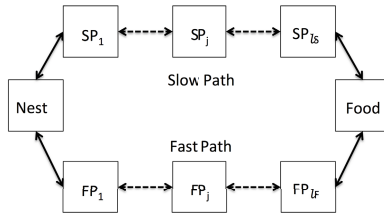


Fig. 4. Schematic of a *Fast* Path and a *Slow* Path for ants; FP_j represents stage j of the *Fast* Path and SP_j represents stage j of the *Slow* Path; $\ell_F < \ell_S$ is assumed

$$\begin{aligned}
\text{Ant@Nest} &= (\text{NFF}, \mathbf{read}(\text{Phe@FP}_1)).\text{AntToFood@FP}_1 + \\
&\quad (\text{NFS}, \mathbf{read}(\text{Phe@SP}_1)).\text{AntToFood@SP}_1 \\
\text{AntToFood@FP}_j &= (\text{NFF}_j, \mathbf{out}(\text{Phe@FP}_j)).\text{AntToFood@FP}_{j+1} & j = 1 \dots \ell_F - 1 \\
\vdots & \\
\text{AntToFood@FP}_{\ell_F} &= (\text{NFF}_{\ell_F}, \mathbf{out}(\text{Phe@FP}_{\ell_F})).\text{Ant@Food} \\
\text{AntToFood@SP}_j &= (\text{NFS}_j, \mathbf{out}(\text{Phe@SP}_j)).\text{AntToFood@SP}_{j+1} & j = 1 \dots \ell_S - 1 \\
\vdots & \\
\text{AntToFood@SP}_{\ell_S} &= (\text{NFS}_{\ell_S}, \mathbf{out}(\text{Phe@SP}_{\ell_S})).\text{Ant@Food} \\
\text{Ant@Food} &= (\text{FNF}_{\ell_F}, \mathbf{read}(\text{Phe@FP}_{\ell_F})).\text{AntToNest@FP}_{\ell_F} + \\
&\quad (\text{FNS}_{\ell_S}, \mathbf{read}(\text{Phe@SP}_{\ell_S})).\text{AntToNest@SP}_{\ell_S} \\
\text{AntToNest@FP}_{\ell_F-j} &= (\text{FNF}_{\ell_F-j}, \mathbf{out}(\text{Phe@FP}_{\ell_F-j})).\text{AntToNest@FP}_{\ell_F-(j+1)} & j = 1 \dots \ell_F - 1 \\
\vdots & \\
\text{AntToNest@FP}_1 &= (\text{FNF}_1, \mathbf{out}(\text{Phe@FP}_1)).\text{Ant@Nest} \\
\text{AntToNest@SP}_{\ell_S-j} &= (\text{FNS}_{\ell_S-j}, \mathbf{out}(\text{Phe@SP}_{\ell_S-j})).\text{AntToNest@SP}_{\ell_S-(j+1)} & j = 1 \dots \ell_S - 1 \\
\vdots & \\
\text{AntToNest@SP}_1 &= (\text{FNS}_1, \mathbf{out}(\text{Phe@SP}_1)).\text{Ant@Nest} \\
\text{ExpPhe@FP}_j &= (\text{PHF}, \mathbf{in}(\text{Phe@FP}_j)).\text{ExpPhe@FP}_j & j = 1 \dots \ell_F \\
\text{ExpPhe@SP}_j &= (\text{PHS}, \mathbf{in}(\text{Phe@SP}_j)).\text{ExpPhe@SP}_j & j = 1 \dots \ell_S
\end{aligned}$$

Fig. 5. ODELINDA model for foraging ants

State Ant@Nest (Ant@Food , respectively) represents an ant at the *Nest* (at the *Food* site, respectively). State AntToFood@FP_j (AntToFood@SP_j , respectively) represents an ant in stage j of the *Fast* (*Slow*, respectively) path, when travelling from the *Nest* to the *Food*. State AntToNest@FP_j (AntToNest@SP_j , respectively) represents an ant in stage j of the *Fast* (*Slow*, respectively) path, when travelling from the *Food* to the *Nest*. For the sake of simplicity, in this model, once an ant leaves the *Nest*, it can only proceed to the *Food* and then come back to the *Nest* (i.e. an ant cannot change its mind half-way a path or get stuck there). This is common in foraging ants models (see [22] and references therein). Finally, processes ExpPhe@FP_j and ExpPhe@SP_j are used for modelling pheromone decay. The definitions of rate functions NFF , NFS , NFF_j , NFS_j , PHF , and PHS are given below, where parameters k , m and p will be discussed later on in this section:

$$\begin{aligned}
\text{NFF} &= \frac{(k + \#\text{Phe@FP}_1)^2}{(k + \#\text{Phe@FP}_1)^2 + (k + \#\text{Phe@SP}_1)^2} \cdot \#\text{Ant@Nest} \\
\text{NFS} &= \frac{(k + \#\text{Phe@SP}_1)^2}{(k + \#\text{Phe@FP}_1)^2 + (k + \#\text{Phe@SP}_1)^2} \cdot \#\text{Ant@Nest} \\
\text{NFF}_j &= m \cdot \#\text{AntToFood@FP}_j \\
\text{NFS}_j &= m \cdot \#\text{AntToFood@SP}_j \\
\text{PHF} &= p \cdot \#\text{ExpPhe@FP}_j \\
\text{PHS} &= p \cdot \#\text{ExpPhe@SP}_j
\end{aligned}$$

The expressions for the definitions of NFF and NFS are written in accordance with results from experimental studies on colonies of Argentine ants, as discussed

in [14,12,22]. The definition of functions FNF , FNS , FNF_j , and FNS_j are similar to those of NFF , NFS , NFF_j and NFS_j due to the symmetry of the model.

In the following we present some analysis results for two specific instantiations of the model and its parameters. We consider a model where the *Fast* path is composed of two stages while the *Slow* path is seven stages long, i.e. $\ell_F = 2$ and $\ell_S = 7$.

We first assume that pheromones do not decay: they accumulate in the path stages so that their total amount grows larger and larger. This is achieved by setting the decay rate p to zero. The value chosen for k is 10, while the ants rate of movement m from one path stage to the next one is set to 0.1. Fig. 6 shows the solution of the equations for the first 500 time units for an initial number of 1000 ants in the *Nest*, while no ants are assumed present in any other path stage, neither at the *Food* site, initially. One unit of pheromone is assumed to be present at time 0 at the stages of the *Slow* and the *Fast* paths closest to the *Nest*. Fig. 6 (left) shows that there is a quick drop in the number of ants at the *Nest* and that for a brief time frame of about 50 time units the cumulative number of ants on the *Slow* path is actually higher than that on the *Fast* path. This situation changes rapidly when ants start to return from the *Food* to the *Nest* providing implicitly feedback to the system by reinforcing the pheromone trace on the *Fast* path. This leads to a rather quick convergence of ants on the *Fast* path. The cumulative amount of pheromones on the *Fast* and the *Slow* path is shown in Fig. 6 (right).

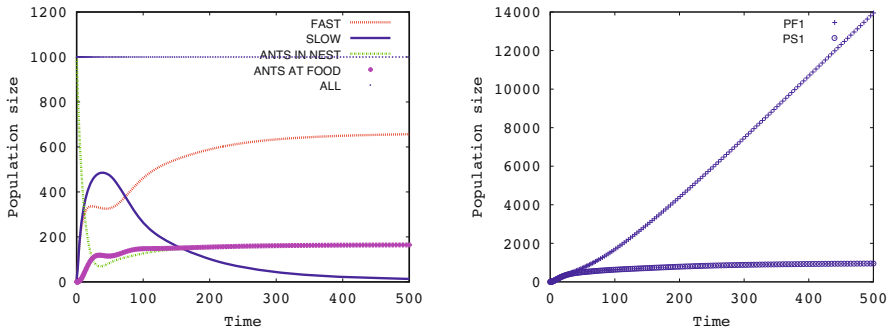


Fig. 6. Evolution over the first 500 time units of the number of ants in the *Nest*, at the *Food* and on the *Fast* and the *Slow* path (left) and the amounts of pheromones on the start of the short (PF1) and long (PS1) path (right)

Fig. 7 shows the results for a variant of the model where pheromone decays with constant rate $p = 0.03$. The evolution of both the cumulative number of ants in various locations and the amount of pheromone on the paths is shown over a time interval of 500 time units. Also in this case the ants converge on the *Fast* path and they are doing so in shorter time than in the case without decay of pheromones. Fig. 7 (left) has been obtained using Octave⁴ for solving

⁴ See for information on Octave <http://www.octave.org>. Version 3.4.0 was used.

the ODE for the model specification. Fig. 7 (right) shows the results obtained via stochastic simulation for the same model with 1000 ants taking the average over 100 runs⁵.

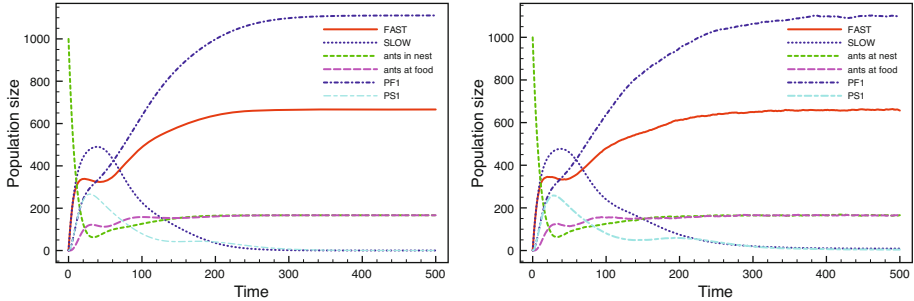


Fig. 7. Evolution over the first 500 time units of the system allowing decay of pheromones with rate 0.03. Solution of the differential equations (left) and stochastic simulation average over 100 runs for a model with 1000 ants (right)

We close this section showing the application of the mean field model-checker FlyFast [19] on the foraging ants example. Fluid model-checking techniques have recently been proposed as scalable techniques for the verification of properties of one (or a few) agents in the context of large populations [4]. These techniques are based on differential semantics, or on difference equations, when considering their discrete time counterparts, as is the case for FlyFast. The input language of the model-checker does not support the specification of models with dynamic population size, but if we can assume sufficiently large upper bounds on the sizes of the data sub-populations for the time horizon of interest⁶, it is rather straightforward to translate the model of foraging ants shown in Fig. 5 into such a language, modelling data by two-state (i.e. “present” and “absent”) processes and using an appropriate scaling of rates to turn the stochastic model into an equivalent probabilistic one [18]. We briefly illustrate the results for two properties for the ants model in Fig. 8. Property A shows how the probability of an ant in the nest to move to the short path within 30 time units changes over time due to the pheromones left behind by other ants. Property B shows the probability to reach a system state within t time units, where t ranges from 0 to 500, in which an ant in the nest moves to the short path within 30 time units with a probability of more than 0.95. Both properties can be expressed using the standard PCTL logic, a probabilistic extension of CTL [16]. Model-checking times for property A is 10 ms, whereas that of property B is 41,047 ms.

The purpose of the foraging ants example is to illustrate the ODELINDA language and mean field analysis approach for asynchronous, tuple based languages.

⁵ Experiments were conducted with a 1.8 GHz Core i7 Intel processor and 4 GB RAM running Mac OS X 10.7.5.

⁶ This can be checked using the result of an ODE analysis or simulation.

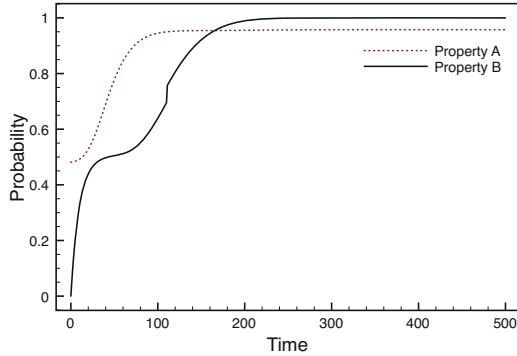


Fig. 8. Mean field model-checking results for properties A and B

Results better matching those of the original experiments in [14,12] can be obtained by a somewhat more complicated model in which ants leave the nest at a constant rate and in which the length of the paths and the constant traversal times are more accurately modelled by adding further path stages on each path implicitly using an Erlang distribution with more stages to approximate the constant traversal times. We omitted this here for the sake of simplicity.

5 Conclusions and Future Work

In this paper we have provided a differential semantics for languages with an asynchronous buffered communication interaction paradigm, e.g. data-repository-/tuple-based ones. In particular, we have defined an individual-based Markovian as well as population based differential semantics for ODELINDA, a simple data-repository-based language. As example of use of the language we have shown a benchmark model of Foraging Ants and some results of its ODE-based analysis. There are several lines of research we plan to follow for moving from a simple experimental kernel language like ODELINDA to a complete, full fledged population modelling language. One line of research focuses of the introduction of an appropriate notion of *space*. One possibility is to take StoKLAIM [11] as a starting point, thus using a simple, locality based approach. Another, perhaps more interesting possibility, instead, is to use a richer, predicate based, addressing mechanism, like (a possibly restricted version of) the addressing mechanism of StocS [20], where the location is just one of the agents' attributes and its values are instances of an appropriate data type, namely *space*. This can take different forms, from topological spaces—including bi- or tri-dimensional continuous space—to more general closure spaces—including generic graphs—as in [6]. Another issue is the inclusion of richer data and operations including templates and pattern-matching. This implies the definition of syntactical restrictions, or static analysis techniques, for guaranteeing that in all computations of a model specification the set of *distinct* data values is bounded, while the multiplicity of each item can of course be unbounded. This also holds for the inclusion of process spawning and processes to be stored/retrieved

to/from repositories. Finally, we plan to adapt fluid model-checking techniques to tuple-based languages.

References

1. Antonaki, M., Philippou, A.: A process calculus for spatially-explicit ecological models. In: Ciobanu, G. (ed.) Proceedings 6th Workshop on Membrane Computing and Biologically Inspired Process Calculi, MeCBIC 2012, Newcastle, UK, vol. 100, pp. 14–28. EPTCS (September 8, 2012), <http://dx.doi.org/10.4204/EPTCS.100.2>
2. Bortolussi, L., Hillston, J., Latella, D., Massink, M.: Continuous approximation of collective systems behaviour: A Tutorial. *Performance Evaluation - An International Journal* 70, 317–349 (2013), doi:10.1016/j.peva.2013.01.001
3. Bortolussi, L., Policriti, A.: Dynamical systems and stochastic programming: To ordinary differential equations and back. *T. Comp. Sys. Biology* 11, 216–267 (2009), doi:10.1007/978-3-642-04186-0_11
4. Bortolussi, L., Hillston, J.: Fluid model checking. In: Koutny, M., Ulidowski, I. (eds.) CONCUR 2012. LNCS, vol. 7454, pp. 333–347. Springer, Heidelberg (2012), http://dx.doi.org/10.1007/978-3-642-32940-1_24
5. Carriero, N., Gelernter, D., Mattson, T.G., Sherman, A.H.: The linda[®] alternative to message-passing systems. *Parallel Computing* 20(4), 633–655 (1994), [http://dx.doi.org/10.1016/0167-8191\(94\)90032-9](http://dx.doi.org/10.1016/0167-8191(94)90032-9)
6. Ciancia, V., Latella, D., Loreti, M., Massink, M.: Specifying and verifying properties of space. In: Diaz, J., Lanese, I., Sangiorgi, D. (eds.) TCS 2014. LNCS, vol. 8705, pp. 222–235. Springer, Heidelberg (2014), http://dx.doi.org/10.1007/978-3-662-44602-7_18
7. Ciocchetta, F., Hillston, J.: Bio-pepa: A framework for the modelling and analysis of biological systems. *Theor. Comput. Sci.* 410(33-34), 3065–3084 (2009), <http://dx.doi.org/10.1016/j.tcs.2009.02.037>
8. De Nicola, R., et al.: The SCEL Language: Design, Implementation, Verification. In: Wirsing, M., Hölzl, M., Koch, N., Mayer, P. (eds.) *Collective Autonomic Systems*. LNCS, vol. 8998, pp. 3–71. Springer, Heidelberg (2015)
9. De Nicola, R., Latella, D., Loreti, M., Massink, M.: A Uniform Definition of Stochastic Process Calculi. *ACM Computing Surveys* 46(1), 5:1–5:35 (2013), doi:10.1145/2522968.2522973
10. De Nicola, R., Ferrari, G.L., Pugliese, R.: KLAIM: A kernel language for agents interaction and mobility. *IEEE Trans. Software Eng.* 24(5), 315–330 (1998), <http://doi.ieeecomputersociety.org/10.1109/32.685256>
11. De Nicola, R., Katoen, J., Latella, D., Loreti, M., Massink, M.: Model checking mobile stochastic logic. *Theor. Comput. Sci.* 382(1), 42–70 (2007), <http://dx.doi.org/10.1016/j.tcs.2007.05.008>
12. Deneubourg, J.L., Aron, S., Goss, S., Pasteels, J.M.: The self-organizing exploratory pattern of the argentine ant. *Journal of Insects Behaviour* 3(2) (1990)
13. Feng, C., Hillston, J.: PALOMA: A Process Algebra for Located Markovian Agents. In: Norman, G., Sanders, W. (eds.) QEST 2014. LNCS, vol. 8657, pp. 265–280. Springer, Heidelberg (2014)
14. Goss, S., Aron, S., Deneubourg, J.L., Pasteels, J.M.: Self-organized shortcuts in the Argentine Ant. *Naturwissenschaften* 76, 579–581 (1989)
15. Guenther, M.C., Bradley, J.T.: Higher moment analysis of a spatial stochastic process algebra. In: Thomas, N. (ed.) EPEW 2011. LNCS, vol. 6977, pp. 87–101. Springer, Heidelberg (2011), http://dx.doi.org/10.1007/978-3-642-24749-1_8

16. Hansson, H., Jonsson, B.: A logic for reasoning about time and reliability. *Formal Aspects of Computing. The International Journal of Formal Methods* 6(5), 512–535 (1994)
17. Hermanns, H., Herzog, U., Katoen, J.: Process algebra for performance evaluation. *Theor. Comput. Sci.* 274(1-2), 43–87 (2002), [http://dx.doi.org/10.1016/S0304-3975\(00\)00305-4](http://dx.doi.org/10.1016/S0304-3975(00)00305-4)
18. Latella, D., Loretì, M., Massink, M.: On-the-fly Fluid Model Checking via Discrete Time Population Models. Extended Version. Technical Report TR-QC-08-2014, QUANTICOL (2014)
19. Latella, D., Loretì, M., Massink, M.: On-the-fly fast mean-field model-checking. In: Abadi, M., Lluch Lafuente, A. (eds.) TGC 2013. LNCS, vol. 8358, pp. 297–314. Springer, Heidelberg (2014), http://dx.doi.org/10.1007/978-3-319-05119-2_17
20. Latella, D., Loretì, M., Massink, M., Senni, V.: Stochastically timed predicate-based communication primitives for autonomic computing. In: Bertrand, N., Bortolussi, L. (eds.) Proceedings Twelfth International Workshop on Quantitative Aspects of Programming Languages and Systems, QAPL 2014, Grenoble, France, April 12-13, vol. 154, pp. 1–16. EPTCS (2014), <http://dx.doi.org/10.4204/EPTCS.154.1>
21. Mamei, M., Zambonelli, F.: Programming pervasive and mobile computing applications: The TOTA approach. *ACM Trans. Softw. Eng. Methodol.* 18(4) (2009), <http://doi.acm.org/10.1145/1538942.1538945>
22. Massink, M., Latella, D.: Fluid analysis of foraging ants. In: Sirjani, M. (ed.) COORDINATION 2012. LNCS, vol. 7274, pp. 152–165. Springer, Heidelberg (2012), http://dx.doi.org/10.1007/978-3-642-30829-1_11
23. Tribastone, M., Gilmore, S., Hillston, J.: Scalable differential analysis of process algebra models. *IEEE Transactions on Software Engineering. IEEE CS* 38(1), 205–219 (2012)