Information and Computation 247 (2016) 37–86

Contents lists available at ScienceDirect

Information and Computation

www.elsevier.com/locate/yinco

Hybrid behaviour of Markov population models

Luca Bortolussi [a,b,*]

[a] Department of Mathematics and Geosciences, University of Trieste, Italy
[b] CNR/ISTI, Pisa, Italy

ARTICLE INFO

Article history:
Received 10 January 2013
Available online 23 December 2015

Keywords:
Stochastic process algebras
Stochastic concurrent constraint programming
Stochastic hybrid systems
Mean field
Fluid approximation
Weak convergence

ABSTRACT

We investigate the behaviour of population models, specified in stochastic Concurrent Constraint Programming (**sCCP**). In particular, we focus on models from which we can define a semantics both in terms of Continuous Time Markov Chains (CTMC) and in terms of Stochastic Hybrid Systems, in which some populations are approximated continuously, while others are kept discrete. We will prove the correctness of the hybrid semantics from the point of view of the limiting behaviour of a sequence of models for increasing population size. More specifically, we prove that, under suitable regularity conditions, the sequence of CTMC constructed from **sCCP** programs for increasing population size converges to the hybrid system constructed by means of the hybrid semantics. We investigate in particular what happens for **sCCP** models in which some transitions are guarded by boolean predicates or in the presence of instantaneous transitions.

1. Introduction

Many complex systems, from genetic networks and signal pathways in cells, to epidemic spreading, to telecommunication systems, can be formally described as Markovian stochastic processes of the class of Markov population models [1]. Stochastic Process Algebras (SPA) are a powerful framework for quantitative modelling and analysis of such population processes [2]. They have been applied in a wide variety of contexts, including computer systems [2], biological systems [3–5], ecological [6] and crowd [7] modelling.

However, their standard semantics, given in terms of Continuous Time Markov Chain (CTMC [8]), suffer from the problem of state space explosion, which impedes the use of SPA to analyze models with a large state space. A recent technique introduced to tackle this problem is fluid approximation [9], which describes the number of system components by means of continuous variables and interprets rates as flows, thus providing a semantics in terms of Ordinary Differential Equations (ODE).

The relationship between these two semantics is grounded on the law of large numbers for population processes [10], first proved by Kurtz back in the seventies [11]. Applying this theoretical framework to SPA models, one obtains that the fluid ODE is the limit of the sequence of CTMC models [12–14], obtained by the standard SPA semantics for increasing system size, usually the total number of agents in the system. This also provides a link with a large body of mathematical literature on fluid and mean field approximation (see e.g. [1] for a recent review).

These results provide the asymptotic correctness of the fluid semantics and justify the use of ODE to analyze large collective SPA models. Fluid approximation is also entering into the analysis phase in a more refined way than just by

* Correspondence to: Department of Mathematics and Geosciences, University of Trieste, Via A. Valerio 12/1, 34127 Trieste, Italy.
E-mail address: luca@dmi.units.it.

http://dx.doi.org/10.1016/j.ic.2015.12.001

numerical simulation. For instance, in [15], the authors use fluid approximation for the computation of passage-times, while in [16,17] the fluid approximation scheme is used to model check properties of single agents in a large population against CSL properties.

Despite the remarkable success of fluid approximation of SPA models, its applicability is restricted to situations in which all components are present in large quantities, and all events depend continuously on the number of the different agent types. This excludes many interesting situations, essentially all those in which some sub-populations have a fixed and small size. This is the case in biological systems, when one considers gene networks, but also in computer systems when one models some form of centralized controller. Furthermore, the description of control policies is often simplified by using forced (or instantaneous) events, happening as soon as certain condition are met, and more generally guard predicates, modulating the set of enabled actions as a function of the global state of the system.

These features of population models are not easily captured in a fluid flow scheme, as they lead naturally to hybrid systems, in which discrete and continuous dynamics coexist. To deal with these situations, in [18,19] the authors proposed a hybrid semantics for a specific SPA, namely stochastic Concurrent Constraint Programming (**sCCP** [4]), associating with a **sCCP** model a hybrid system where continuous dynamics is interleaved with discrete Markovian jumps. In [20], also instantaneous transitions are incorporated in the framework. Such a hybrid semantics has been also defined for other process algebras, namely PEPA [21] and Bio-PEPA [22]. In this way, one can circumvent the limits of fluid-flow approximation, whilst keeping discrete only the portions of the system that cannot be safely described as continuous. Roughly speaking, this hybrid semantics works by first identifying a subset of system variables to be approximated continuously, keeping discrete the remaining ones. The latter set of variables identifies the discrete skeleton of the hybrid system, while the former defines the continuous state space. Then, each activity of agents, corresponding to a transition that modifies the state of the system, is classified as continuous, discrete/stochastic, or discrete/instantaneous. The first class of transitions is used to construct a vector field giving the continuous dynamics of the hybrid system (in each mode), while the other two transition classes define the discrete dynamics. It is worth stressing that such a hybrid approximation is generally applicable to population models, SPA offering however a nice formal framework to automatically construct such hybrid limits.

The advantages of working with a hybrid semantics for SPA are mainly rooted in the speed-up that can be achieved in the simulation, as discussed e.g. in [19] and [23]. Moreover, the hybrid semantics put at disposal of the modeller a broader set of analysis tools, like transient computation [24] or moment closure techniques [25,26].

While the theory of deterministic approximation of CTMC is well developed, hybrid approximation has attracted much less attention. To the author's knowledge, the preliminary work [27] on which this paper is based was the first attempt to prove hybrid convergence results in a formal method setting. There has been some previous work on hybrid limits in [28], restricted however to a specific biological example, and in the context of large deviation theory [29], where deterministic approximation of models with level variables has been considered (but in this case transitions between mODE are fast, so that the discrete dynamics is always at equilibrium in the limit). More recent work is [30], which discusses hybrid limits for genetic networks (essentially the class of models considered in [27] with some extensions).

The focus of this paper is to provide a general framework to infer consistence of hybrid semantics of SPA models in the light of asymptotic correctness. In doing this, we aimed for generality, proving hybrid limit theorems for a framework including instantaneous events, with guards possibly involving model time, random resets, and guards in continuous and stochastic transitions. The goal was to identify a broad set of conditions under which convergence holds, potentially usable in static analysis algorithmic procedures that check if a given model satisfies the conditions for convergence. We will comment on this issue in several points in the paper. To author's knowledge, this is the first attempt to discuss hybrid approximation in such generality.

We will start our presentation recalling **sCCP** (Section 2.1) and the hybrid semantics (Section 2.3). We will formally define it in terms of Piecewise Deterministic Markov Processes (Section 2.4, [31]), a class of Stochastic Hybrid Processes in which the continuous dynamics is given in terms of Ordinary Differential Equations, while the discrete dynamics is given by forced transitions (firing as soon as their guard becomes true) and by Markovian jumps, firing with state dependent rate. The hybrid semantics is defined by introducing an intermediate layer in terms of an automata based description, by the so-called Transition-Driven Stochastic Hybrid Automata (TDSHA, Sections 2.2 and 2.5, [18,19]).

After presenting the classic fluid approximation result, recast in our framework (Section 4), we turn our attention to **sCCP** models that are converted to TDSHA containing only discrete/stochastic and continuous transitions, with no guards and no instantaneous transitions, but allowing random resets (general for discrete/stochastic transitions and restricted for continuous ones). In Section 5, we prove a limit theorem under mild consistency conditions on rates and resets, showing that the sequence of CTMC associated with a **sCCP** program, for increasing system size, converges to the limit PDMP in the sense of weak convergence. Technically speaking, the appearance of weak convergence instead of convergence in probability, in which classic fluid limit theorems are usually stated, depends on the fact that the limit process is stochastic and can have discontinuous trajectories.

We then turn our attention to the limit behaviour in the presence of sources of discontinuity, namely instantaneous transitions (Section 6) or guards in continuous (Sections 7.2 and 7.1) or discrete/stochastic transitions (Section 7.4).

In all these cases, the situation is more delicate and the conditions for convergence are more complex. Guards in continuous transitions introduce discontinuities in the limit vector fields, requiring us to define the continuous dynamics in terms of the so-called piecewise-smooth dynamical systems [32] or, more generally, in terms of differential inclusions [33]. Here,