Theoretical Computer Science 563 (2015) 99–143

Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs

On the tradeoff between compositionality and exactness in weak bisimilarity for integrated-time Markovian process calculi

Marco Bernardo

Dipartimento di Scienze di Base e Fondamenti, Università di Urbino, Italy

ARTICLE INFO

Article history:
Received 16 October 2012
Received in revised form 28 May 2014
Accepted 15 October 2014
Available online 5 November 2014
Communicated by G. Rozenberg

Keywords:
Stochastic process algebra
Weak bisimulation equivalence
Compositionality
Continuous-time Markov chains
Pseudo-aggregations
Exactness

ABSTRACT

Integrated-time Markovian process calculi rely on actions whose durations are quantified by exponentially distributed random variables. The Markovian bisimulation equivalences defined so far for these calculi treat exponentially timed internal actions like all the other actions, because each such action has a nonzero duration and hence can be observed if it is executed between a pair of exponentially timed noninternal actions. However, no difference may be noted, at stationary state, between a sequence of exponentially timed internal actions and a single exponentially timed internal action, if their expected durations and execution probabilities coincide, a fact exploited in Hillston's weak isomorphism. We show that Milner's approach can be adapted on the basis of this fact, so to derive a weak bisimulation equivalence for integrated-time Markovian process calculi, up to a tradeoff between compositionality and exactness inherent to the Markovian setting. The resulting weak Markovian bisimulation equivalence induces a pseudo-aggregation that is exact at stationary state for all the considered processes, but turns out to be a congruence only over sequential processes. To achieve compositionality over concurrent processes, we need to enhance the abstraction capability of the equivalence in the presence of interleaved computations. However, the corresponding pseudo-aggregation turns out to be exact at stationary state only for a subset of concurrent processes. In addition to this tradeoff, we present, for the first equivalence, a sound and complete axiomatization over sequential processes, which is instrumental to characterize pseudo-aggregations, and a polynomial-time equivalence-checking algorithm, which can be exploited for the compositional minimization of concurrent processes.

1. Introduction

Quantitative models based on continuous-time Markov chains (see, e.g., [34]) like stochastic Petri nets (see, e.g., [2]) and stochastic process algebras (see, e.g., [22,20]) have been deeply investigated and successfully used in the last decades to predict the performance of computer, communication, and software systems. From a conceptual viewpoint, we can distinguish between *integrated-time* and *orthogonal-time* Markovian models [9]. In the former, which are more natural for modeling purposes, the passage of time is associated with the execution of activities, i.e., activities are considered *durational*. In the

E-mail address: marco.bernardo@uniurb.it.

http://dx.doi.org/10.1016/j.tcs.2014.10.025

latter, which are more elegant on the theoretical side, the passage of time is separate from the execution of activities, i.e., activities are *durationless* and hence time passing has to be represented explicitly.

Several Markovian behavioral equivalences (see [3] and the references therein) have been proposed in the literature for relating and manipulating system models with an underlying continuous-time Markov chain (CTMC) semantics. These equivalences are extensions of the traditional approaches to the definition of behavioral equivalences, and take into account time passing described by means of exponential distributions. A feature shared by relations like Markovian bisimilarity, Markovian testing equivalence, and Markovian trace equivalence is that of being *strong*, in the sense that they treat internal activities – which cannot be seen by an external observer – like the other activities. Only a few variants investigated in [20, 31,25,12] are able to abstract from internal activities and/or purely probabilistic branchings.

The useful capability of abstracting from internal actions can be easily achieved in the orthogonal-time setting, because in that case activities are immediate (i.e., take no time) and hence well-known techniques developed for nondeterministic processes can be employed to get rid of these activities when they are internal. Let us denote by τ the invisible or silent action. In the nondeterministic setting, a process that can perform action a followed by action τ and action b and then terminates – written $a\,.\,\tau\,.\,b\,.\,\underline{0}$ – is weakly equivalent to a process that can perform action a followed by action b and then terminates – written $a\,.\,b\,.\,\underline{0}$. The situation is more complicated in the integrated-time setting. Since actions have exponentially distributed durations – uniquely identified by positive real numbers called rates – it is not necessarily the case that simplifications like the one above can be made.

For instance, if action a has rate λ, action b has rate μ, and action τ has rate γ, the two resulting integrated-time Markovian processes $\langle a, \lambda\rangle.\langle \tau, \gamma\rangle.\langle b, \mu\rangle.\underline{0}$ and $\langle a, \lambda\rangle.\langle b, \mu\rangle.\underline{0}$ are not weakly equivalent. In fact, recalling that the expected duration of an action coincides with the reciprocal of the rate of the action, the former process has a maximal computation whose expected duration is $\frac{1}{\lambda} + \frac{1}{\gamma} + \frac{1}{\mu}$, whereas the latter process has a maximal computation whose expected duration is $\frac{1}{\lambda} + \frac{1}{\mu}$. From another viewpoint, in the former case an external observer would see an a-action for an amount of time t_λ and a b-action for an amount of time t_μ, with a delay t_γ in between, while in the latter case the external observer would not see any delay between the termination of the execution of a and the beginning of the execution of b. Therefore, in a Markovian setting, a τ-action executed between a pair of non-τ-actions cannot be abstracted away, because it has a nonzero duration and hence can be, from a timing viewpoint, observed.

Hillston's weak isomorphism [22] indicates that we should not be too pessimistic. As a different example, take a process that, between actions a and b, can perform *two* τ-actions with rates γ_1 and γ_2, respectively: $\langle a, \lambda\rangle.\langle \tau, \gamma_1\rangle.\langle \tau, \gamma_2\rangle.\langle b, \mu\rangle.\underline{0}$. In this case, an observer may not be able to distinguish between the execution of the two τ-actions above and the execution of *a single τ-action whose expected duration is the sum of the expected durations of the two original τ-actions*, i.e., $\frac{1}{\gamma_1} + \frac{1}{\gamma_2} = \frac{\gamma_1+\gamma_2}{\gamma_1\cdot\gamma_2}$. In other words, the process may be viewed as being *weakly equivalent* to $\langle a, \lambda\rangle.\langle \tau, \frac{\gamma_1\cdot\gamma_2}{\gamma_1+\gamma_2}\rangle.\langle b, \mu\rangle.\underline{0}$.

The two processes above are certainly weakly equivalent from a functional standpoint. However, since the sum of the two exponential random variables quantifying the durations of the two original τ-actions has been *approximated* with a single mean-preserving exponential random variable, it is not necessarily the case that the two processes have the same performance characteristics. This would be true if the equivalence induced a *pseudo*-aggregation of the underlying CTMC that is exact, i.e., such that the transient/stationary probability of being in a macrostate of the aggregated stochastic process – which is *assumed* to be a CTMC – is the sum of the transient/stationary probabilities of being in one of the constituent microstates of the original CTMC. This is the case with Markovian bisimilarity, which is in agreement with the well-known exact CTMC-level aggregation called ordinary lumpability [22,16], and Markovian testing and trace equivalences, which are consistent with a coarser exact CTMC-level aggregation called T-lumpability [8,33].

In this paper, we show that the construction used in [27] to derive a weak bisimulation equivalence for nondeterministic process calculi can be extended to integrated-time Markovian process calculi. The resulting equivalence is weak in the sense that it is capable of *abstracting from the number and the order of consecutive exponentially timed τ-actions* in a computation. It reduces any such sequence to a *single* exponentially timed τ-action *preserving both the expected duration and the execution probability of the original action sequence*. From a stochastic viewpoint, this reduction amounts to replacing hypoexponentially distributed durations with exponentially distributed durations having the same expected value. As a consequence, processes related by the resulting equivalence will not possess the same transient performance measures, unless they refer to properties expressed as the *mean time to certain events*. However, those processes may possess the same stationary reward-based performance measures, as the pseudo-aggregation induced by the considered equivalence on the CTMC underlying each process may be *exact at stationary state*.

Defining a weak Markovian bisimilarity that works as outlined above causes a *tradeoff between semantical compositionality and pseudo-aggregation exactness* to emerge, which is inherent to the Markovian setting. For this reason, we divide the presentation of our results into two parts.

Firstly, we extend the construction of [27] in the simplest possible way, so that the only sequences of exponentially timed internal transitions that are reduced are those that traverse *states enabling only exponentially timed internal actions*. The resulting weak Markovian bisimulation equivalence induces a pseudo-aggregation – called W-lumpability – that is exact at stationary state for *all* the considered processes, thus ensuring full preservation of stationary reward-based performance measures. However, the equivalence is a congruence *only over sequential processes*, a fact that limits its usefulness for state space minimization purposes when there are several processes composed in parallel.