Aalborg Universitet

A faster-than relation for semi-Markov decision processes

Pedersen, Mathias Ruggaard; Bacci, Giorgio; Larsen, Kim Guldstrand

[Link to publication from Aalborg University](#)

A Faster-Than Relation for Semi-Markov Decision Processes

Mathias Ruggaard Pedersen   Giorgio Bacci   Kim Guldstrand Larsen

Department of Computer Science
Aalborg University
Denmark

{mrp,grbacci,kgl}@cs.aau.dk

When modeling concurrent or cyber-physical systems, non-functional requirements such as time are important to consider. In order to improve the timing aspects of a model, it is necessary to have some notion of what it means for a process to be faster than another, which can guide the stepwise refinement of the model. To this end we study a *faster-than relation* for semi-Markov decision processes and compare it to standard notions for relating systems. We consider the compositional aspects of this relation, and show that the faster-than relation is not a precongruence with respect to parallel composition, hence giving rise to so-called parallel timing anomalies. We take the first steps toward understanding this problem by identifying decidable conditions sufficient to avoid parallel timing anomalies in the absence of non-determinism.

1 Introduction

Timing aspects are important when considering real-time or cyber-physical systems. For example, they are of interest in real-time embedded systems when one wants to verify the worst-case execution time for guaranteeing minimal system performance or in safety-critical systems when one needs to ensure that unavoidable rigid deadlines will always be met [13].

Semi-Markov decision processes are continuous-time Markov decision processes where the residence-time on states is governed by generic distributions on the positive real line. These systems have been extensively used to model real-time cyber-physical systems [17, 27].

For reasoning about timing aspects it is important to understand what it formally means for a real-time or cyber-physical system to operate faster than another. To this end we consider a notion of *faster-than relation* for semi-Markov decision processes. The definition of faster-than relation we propose in this paper is a reactive version of an analogous notion of faster-than relation previously introduced in [21] for the case of generative systems. According to this relation, a semi-Markov decision process is faster than another one when it reacts to any sequence of inputs with equal or higher probability than the slower process, within the same time bound.

Often, complex cyber-physical systems are organised as concurrent systems of multiple components running in parallel and interacting with each other. Such systems are better analysed *compositionally*, that is, by breaking them into smaller components that are more easily examined [5]. However, it is not always the case that an analysis on the components carries over to the full composite system. A well known example of this, occurring in real-time systems such as scheduling for processors [4, 14], are *timing anomalies*, that is, when locally faster behaviour leads to a globally slower behaviour [12].

In this paper we study the compositional aspects of the faster-than relation for semi-Markov decision processes. We are interested in the situation depicted in Figure 1 where we have a composite system consisting of a context W and a component V, and we want to understand what happens when we replace V with another component U that is faster than V. We consider common notions of parallel composition,
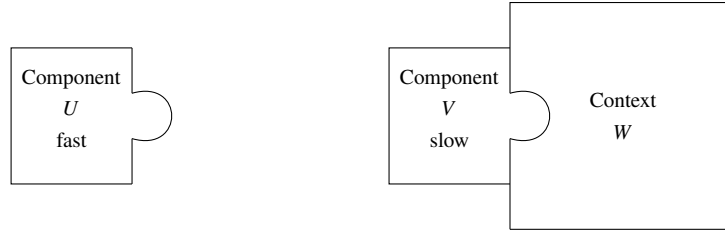
Figure 1: The context W operates in parallel with the component V. If the component U is faster than V, then if we replace V with U, we would expect the overall behaviour to also be faster.

and show that timing anomalies can occur using our faster-than relation, even in the absence of non-determinism. This shows that timing anomalies are not caused by non-determinism, but arise from the linear timing behaviour of processes.

As a first step toward recovering the compositional reasoning for the faster-than relation, we identify a condition, called *monotonicity*, to be checked on the components taking part to the parallel composition which guarantees that timing anomalies cannot occur. Our condition is parametric in the parallel composition operator and applies to all the notions of parallel composition considered in this paper.

Presently, we do not know whether the monotonicity condition is decidable or not, however we introduce another condition, called *strong monotonicity*, still parametric on the parallel composition operator, that we show to be decidable. This decidability result is remarkable for a series of reasons. In [21], the faster-than relation over generic semi-Markov processes was proven to be undecidable and *Positivity-hard* if restricted to processes without non-determinism. The Positivity-hardness relates the faster-than problem to the Skolem problem, an important problem in number theory, whose decidability status has been an open for at least 80 years [20, 1]. In the light of these results, it is not surprising that strong monotonicity condition only applies to processes with no non-determinism.

Related Work. The notion of a faster-than relation has been studied in many different contexts throughout the literature. The work most closely related to ours is that of Pedersen et al. [21], which considers a generative version of the faster-than relation, whereas we study the reactive version. The focus of [21] is on decidability issues, and the faster-than relation is proved undecidable, and the proof can easily be extended to also show that the reactive version is undecidable. However, positive results are also given in the form of an approximation algorithm, and a decidability result for unambiguous processes. Baier et al. [2] define, among other relations, a simulation relation for continuous-time Markov chains which can be interpreted as a faster-than relation, and study its logical characterisation. However, none of these works consider compositional aspects.

For process algebras, discrete-time faster-than relations have been defined for variations of Milner's CCS, and shown to be precongruences with respect to parallel composition [6, 15, 18, 24]. Lüttgen and Vogler [16] attempt to unify some of these process algebraic approaches and also consider the issue of parallel timing anomalies. For Petri nets, Vogler [29, 28] considers a testing preorder as a faster-than relation and shows that this is a precongruence with respect to parallel composition. Geilen et al. [7] introduces a refinement principle for timed actor interfaces under the slogan "the earlier, the better", which can also be seen as an example of a faster-than relation.

Work on timing anomalies date back to at least 1969 [9], but the most influential paper in the area is probably that of Lundqvist and Stenström [14], which shows that timing anomalies can occur in dynamically scheduled processors. More recent work has focused on compositional aspects [12] and defining timing anomalies formally, using transition systems as the formalism [4, 23].

2   Notation and Preliminaries

In this section we fix some notation and recall concepts that are used throughout the rest of the paper. Let $\mathbb{N}$ denote the natural numbers and $\mathbb{R}_{\geq 0}$ denote the non-negative real numbers, which we equip with the standard Borel σ-algebra $\mathbb{B}$. For any set X, $\mathscr{D}(X)$ denotes the set of probability measures on X, and $\mathscr{D}_{\leq}(X)$ the set of subprobability measures on X. For an element $x \in X$, we will use δ_x to denote the Dirac measure at x defined as $\delta_x(E) = 1$ if $x \in E$ and $\delta_x(E) = 0$ otherwise. We fix a non-empty, countable set L of *labels* or *actions* and equip them with the discrete σ-algebra Σ_L.

For a probability measure $\mu \in \mathscr{D}(\mathbb{R}_{\geq 0})$, we denote by F_μ its *cumulative distribution function (CDF)* defined as $F_\mu(t) = \mu([0,t])$, for all $t \in \mathbb{R}_{\geq 0}$. We will denote by $Exp[\theta]$ the CDF of an exponential distribution with rate $\theta > 0$. The *convolution* of two probability measures $\mu, \nu \in \mathscr{D}(\mathbb{R}_{\geq 0})$, denoted by $\mu * \nu$, is the probability measure on $\mathbb{R}_{\geq 0}$ given by $(\mu * \nu)(B) = \int_{-\infty}^{\infty} \nu(B - x)\,\mu(\mathrm{d}x)$, for all $B \in \mathbb{B}$ [3]. Convolution is associative, i.e., $\mu * (\nu * \eta) = (\mu * \nu) * \eta$, and commutative, i.e., $\mu * \nu = \nu * \mu$.

3   Semi-Markov Decision Processes

In this section we recall the definition of semi-Markov decision processes.

Definition 3.1. A *semi-Markov decision process (SMDP)* is a tuple $M = (S, s_0, \tau, \rho)$ where

- S is a non-empty, countable set of *states*,
- $s_0 \in S$ is a distinguished *start state*,
- $\tau : S \times L \to \mathscr{D}_{\leq}(S)$ is a *transition probability function*, and
- $\rho : S \to \mathscr{D}(\mathbb{R}_{\geq 0})$ is a *residence-time probability function*.   ▲

The operational behaviour of an SMDP $M = (S, s_0, \tau, \rho)$ is as follows. Starting from s_0, the process reacts to an external input $a \in L$ provided by the environment by changing its state to $s' \in S$ within time $t \in \mathbb{R}_{\geq 0}$ with probability $\tau(s_0, a)(s') \cdot \rho(s_0)([0,t])$. The process then continues with the new state s' in place of s_0.

Notice that Markov decision processes are a special case of SMDPs where for all $s \in S$, $\rho(s) = \delta_0$ (i.e. transitions happen instantaneously), and that continuous-time Markov decision processes are also a special case of SMDPs where, for all states $s \in S$, $F_{\rho(s)} = Exp[\theta_s]$ for some rate $\theta_s \in \mathbb{R}_{\geq 0}$.

The executions of an SMDP $M = (S, s_0, \tau, \rho)$ are infinite timed transition sequences of the form

$$\pi = (a_1, t_1, s_1)(a_2, t_2, s_2) \cdots \in (L \times \mathbb{R}_{\geq 0} \times S)^{\omega},$$

representing the fact that M waited t_i time units after the action a_i was input and then moved to the state s_i. We will refer to executions of an SMDP as *paths*. We denote by $(\Pi(M), \Sigma)$ the *measurable space of paths*, where Σ is the smallest σ-algebra generated by the cylinders of the form $L_1 \times R_1 \times S_0, \ldots, L_n \times R_n \times S_n$ for $L_i \in 2^L$, $R_i \in \mathbb{B}$, and $S_i \in 2^S$. Given a path π and $i \in \mathbb{N}$, we let $\pi|_i = (a_1, t_1, s_1) \ldots (a_i, t_i, s_i)$ be the prefix up to i of π. We let $\Pi(M)$ denote the set of all timed action paths in M, and denote by

$$\Pi_n(M) = \{\pi|_i \mid \pi \in \Pi(M)\}$$

the set of all prefixes of length n.

In this paper we assume that external choices are resolved by means of memoryless stochastic schedulers, however all the results we present still hold for schedulers that can remember actions and states.

Definition 3.2. Given an SMDP $M = (S, s_0, \tau, \rho)$, a *scheduler* for M is a function $\sigma : S \to \mathscr{D}(L)$ that assigns to each state a probability distribution over action labels.   ▲

We will use the notation $\tau^\sigma(s, a)(s')$ as shorthand for $\tau(s, a)(s') \cdot \sigma(s)(a)$ to denote the probability of moving from state s to s' under the stochastic choice of a given by σ.

Next we recall the standard construction of the measurable space of paths. A *cylinder set* of rank $n \geq 1$ is the set of all paths whose n-th prefix is contained in a common subset $E \subseteq \Pi_n(M)$, and is given by

$$\mathfrak{C}(E) = \{\pi \in \Pi(M) \mid \pi|_n \in E\}.$$

It will be convenient to denote *rectangular cylinders* of the form

$$\mathfrak{C}(L_1 \times R_1 \times S_1 \times \cdots \times L_n \times R_n \times S_n),$$

for $L_i \subseteq L$, $R_i \subseteq \mathbb{R}_{\geq 0}$, and $S_i \subseteq S$ as

$$\mathfrak{C}(L_1 \ldots L_n, R_1 \ldots R_n, S_1 \ldots S_n).$$

We denote by $(\Pi(M), \Sigma)$ the *measurable space of paths*, where Σ is the smallest σ-algebra generated by the measurable cylinders of the form

$$\mathfrak{C}(L_1 \ldots L_n, R_1 \ldots R_n, S_1 \ldots S_n)$$

for $L_i \in 2^L$, $R_i \in \mathbb{B}$, $S_i \in 2^S$. Given an SMDP M and a scheduler σ for it, we define inductively the usual probability on cylinders as follows.

Definition 3.3. Let $M = (S, s_0, \tau, \rho)$ be an SMDP. Given a scheduler σ for M and a state $s \in S$, $\mathbb{P}_M^\sigma(s)$ is defined as the unique (sub)probability measure[1] on $(\Pi(M), \Sigma)$ such that for all $L_i \in 2^L$, $R_i \in \mathbb{B}$, and $S_i \in 2^S$ with $1 \leq i \leq n$, we have

$$\mathbb{P}_M^\sigma(s)(\mathfrak{C}(L_1, R_1, S_1)) = \rho(s)(R_1) \cdot \sum_{a \in L_1} \sum_{s' \in S_1} \tau^\sigma(s, a)(s')$$

and

$$\mathbb{P}_M^\sigma(s)(\mathfrak{C}(L_1 \ldots L_n, R_1 \ldots R_n, S_1 \ldots S_n))$$
$$= \rho(s)(R_1) \cdot \sum_{a \in L_1} \sum_{s' \in S_1} \tau^\sigma(s, a)(s') \cdot \mathbb{P}_M^\sigma(s')(\mathfrak{C}(L_2 \ldots L_n, R_2 \ldots R_n, S_2 \ldots S_n)).   ▲$$

For example, $\mathbb{P}_M^\sigma(s)(\{a\}\{b\}, [0, x][0, y], \{s'\}\{s''\})$ is the probability of doing an a-transition from s to s' within x time units and then doing a b-transition from s' to s'' within y time units.

4 A Faster-Than Relation

Intuitively, a process U is faster than V if it is able to execute any sequence of actions $a_1, \ldots, a_n$ in less time than V. Since our systems are probabilistic, we need to consider the probability of their execution within any time bound.

Consider the two simple SMDPs U and V in Figure 2 with just a single transition label and initial states u_0 and v_0, respectively. Here μ, ν, η are arbitrary probability measures on $\mathbb{R}_{\geq 0}$, representing the

---

[1]Existence and uniqueness is guaranteed by the Hahn-Kolmogorov theorem [26].
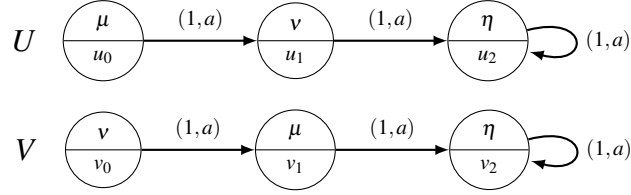
Figure 2: If $F_\mu(t) \geq F_v(t)$ for all t, then U is faster than V in the first states, and after that their probabilities are the same, so U is faster than V.

residence-time distributions at each state. An arrow with label (p,a) means that when a is chosen as the action, then the SMDP takes the transition given by the arrow with probability p. The only finite sequences of actions that can be executed in these SMDPs are of the form a^n for $n > 0$.

For U to be faster than V, it should be the case that for any time t and any scheduler σ for V, we must be able to find a scheduler σ' for U which allows U to execute any sequence of actions within time t with higher or equal probability than V. Hence, the type of events on which we want to focus are the following.

Definition 4.1. For any finite sequence of actions $a_1, \ldots, a_n$, and $t \in \mathbb{R}_{\geq 0}$, we say that

$$\mathfrak{C}(a_1 \ldots a_n, t) = \{(b_1, t_1, s_1)(b_2, t_2, s_2) \cdots \in \Pi(M) \mid \forall 1 \leq i \leq n,\ b_i = a_i \text{ and } \sum_{j=1}^n t_i \leq t\}$$

is a *time-bounded cylinder*. The *length* of a time-bounded cylinder is the length of the sequence of actions in the cylinder. ▲

Note that $\mathfrak{C}(a_1 \ldots a_n, t)$ is measurable in $(\Pi(M), \Sigma)$, since

$$f : \Pi_n(M) \to L^n \times \mathbb{R}_{\geq 0}^n \times S^n$$

given by

$$f((o_1, t_1, s_1), \ldots, (o_n, t_n, s_n)) = ((o_1, \ldots, o_n), (t_1, \ldots, t_n), (s_1, \ldots, s_n))$$

and

$$\mathtt{res}_n : \Pi(M) \to \Pi_n(M)$$

given by

$$\mathtt{res}_n(\pi) = \pi|_n$$

are both measurable, and hence

$$(f \circ \mathtt{res}_n)^{-1}(S^n \times \{(a_1, \ldots, a_n)\} \times B_t^n) = \mathfrak{C}(a_1 \ldots a_n, t)$$

is measurable, where $B_t^n = \{(r_1, \ldots, r_n) \in \mathbb{R}_{\geq 0}^n \mid \sum_{i=1}^n r_i \leq t\}$.

Example 4.2. The time-bounded cylinder $\mathfrak{C}(aa, 2)$ denotes the set of all paths where the first two output labels are both a's, and the first two steps of the path are completed within 2 time units. ♦

For the rest of the paper, we fix three SMDPs $M = (S, s_0, \tau, \rho)$, $U = (S_U, u_0, \tau_U, \rho_U)$, and $V = (S_V, v_0, \tau_V, \rho_V)$. Now we are ready to define what it means for an SMDP to be "faster than" another one.

Definition 4.3 (Faster-than). We say that U is *faster than* V, written $U \preceq V$, if for all schedulers σ for V, time bounds t, and sequences of actions $a_1 \ldots a_n$, there exists a scheduler σ' for U and time bound $t' \leq t$, such that $\mathbb{P}^{\sigma'}(u_0)(\mathfrak{C}(a_1 \ldots a_n, t')) \geq \mathbb{P}^\sigma(v_0)(\mathfrak{C}(a_1 \ldots a_n, t))$. ▲

Clearly, the faster-than relation $\preceq$ is a preorder. The following proposition gives a characterisation of the faster-than relation that is often easier to work with.

Proposition 4.4. *$U \preceq V$ if and only if for all schedulers σ for V there exists a scheduler σ' for U such that $\mathbb{P}^{\sigma'}(u_0)(C) \geq \mathbb{P}^{\sigma}(v_0)(C)$, for all time-bounded cylinders C.*

Before showing an example of an SMDP being faster than another one, we provide an analytic solution for computing the probability over time-bounded cylinders in terms of convolutions of the residence time distributions.

Proposition 4.5. *For any scheduler σ for M, and $s \in S$ we have*

$$\mathbb{P}^{\sigma}(s)(\mathfrak{C}(a_1 \ldots a_n, t)) = \sum_{s_1, \ldots, s_n \in S} \tau^{\sigma}(s, a_1)(s_1) \cdots \tau^{\sigma}(s_{n-1}, a_n)(s_n) \cdot (\rho(s) * \rho(s_1) * \cdots * \rho(s_{n-1}))([0, t]).$$

Proposition 4.5 intuitively says that the absorption-time of any path of length n through the SMDP is distributed as the n-fold convolution of its residence-time probabilities. Therefore, the probability of doing transitions with labels $a_1, \ldots, a_n$ within time t is the sum of the probabilities of taking a path of length n with labels $a_1, \ldots, a_n$ through the SMDP, weighted by the probability of reaching the end of each of these paths within time t. This is similar in spirit to a result on phase-type distributions, see e.g. [22, Proposition 2.11].

From Proposition 4.5 we can also derive the following which gives a more direct inductive definition of the probability on time-bounded cylinders. If we fix $a_1 \ldots a_n$ and let t vary, we get a CDF

$$\mathbb{P}^{\sigma}(s)(a_1 \ldots a_n)([0, t]) = \mathbb{P}^{\sigma}(s)(\mathfrak{C}(a_1 \ldots a_n, t)).$$

Proposition 4.6. *The CDF $\mathbb{P}^{\sigma}(s)(a_1 \ldots a_n)$ can be characterised inductively by*

$$\mathbb{P}^{\sigma}(s)(a_1)([0, t]) = \sum_{s' \in S} \tau^{\sigma}(s, a_1)(s') \cdot \rho(s)([0, t]),$$

$$\mathbb{P}^{\sigma}(s)(a_1 \ldots a_n)([0, t]) = \sum_{s' \in S} \tau^{\sigma}(s, a_1)(s') \cdot (\rho(s) * \mathbb{P}^{\sigma}(s')(a_2 \ldots a_n))([0, t]).$$

Proposition 4.6 also shows that our definition of faster-than coincides with the one from [21], except ours is reactive rather than generative.

Example 4.7. Consider the SMDPs U and V that are depicted in Figure 2. Assuming that $F_\mu(t) \geq F_\nu(t)$ for all t, we now show that $U \preceq V$. To compare U and V, first notice that we only need to consider time-bounded cylinders of the form $\mathfrak{C}(a^n, t)$, for $n \geq 1$. Since the set of actions is $L = \{a\}$, the only possible valid scheduler σ for both U and V is the one assigning the Dirac measure δ_a to all states. We consider two cases.

(Case $n = 1$) In this case we get

$$\mathbb{P}^{\sigma}(u_0)(\mathfrak{C}(a, t)) = F_\mu(t) \quad \text{and} \quad \mathbb{P}^{\sigma}(v_0)(\mathfrak{C}(a, t)) = F_\nu(t).$$

Since we assumed $F_\mu(t) \geq F_\nu(t)$ for all t, this implies

$$\mathbb{P}^{\sigma}(u_0)(\mathfrak{C}(a, t)) \geq \mathbb{P}^{\sigma}(v_0)(\mathfrak{C}(a, t)).$$

(Case $n > 1$) By Proposition 4.5 we have both

$$\mathbb{P}^{\sigma}(u_0)(\mathfrak{C}(a^n, t)) = (\mu * \nu * \eta^{*(n-2)})([0, t]) \quad \text{and} \quad \mathbb{P}^{\sigma}(v_0)(\mathfrak{C}(a^n, t)) = (\nu * \mu * \eta^{*(n-2)})([0, t]),$$

where η^{*n} is the n-fold convolution of η, defined inductively by $\eta^{*0} = \delta_0$ and $\eta^{*(n+1)} = \eta * \eta^{*n}$. Since convolution is commutative and associative, and δ_0 is the identity for convolution, we obtain

$$\mathbb{P}^{\sigma}(u_0)(\mathfrak{C}(a^n, t)) = \mathbb{P}^{\sigma}(v_0)(\mathfrak{C}(a^n, t)).$$

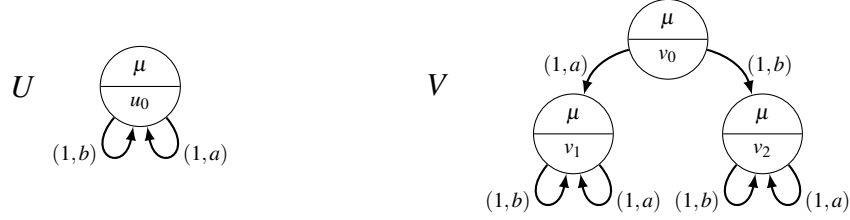We therefore conclude that $U \preceq V$.   ♦

Figure 3: Example showing that the faster-than relation and the simulation relation are incomparable.

4.1 Comparison with Simulation and Bisimulation

The standard notions used to compare processes are bisimulation [19] and simulation [2]. Next we recall their definitions, naturally extended to our setting of SMDPs.

Definition 4.8. For an SMDP M, a relation $R \subseteq S \times S$ is a *bisimulation relation* (resp. *simulation relation*) on M if for all $(s_1, s_2) \in R$ we have

- $F_{\rho(s_1)}(t) = F_{\rho(s_2)}(t)$ (resp. $F_{\rho(s_1)}(t) \leq F_{\rho(s_2)}(t)$) for all $t \in \mathbb{R}_{\geq 0}$ and

- for all $a \in L$ there exists a weight function $\Delta_a : S \times S \to [0,1]$ such that

  - $\Delta_a(s, s') > 0$ implies $(s, s') \in R$,
  - $\tau(s_1, a)(s) = \sum_{s' \in S} \Delta_a(s, s')$ for all $s \in S$, and
  - $\tau(s_2, a)(s') = \sum_{s \in S} \Delta_a(s, s')$ for all $s' \in S$.

If there is a bisimulation relation (resp. simulation relation) R such that $(s_1, s_2) \in R$, then we say that s_1 and s_2 are *bisimilar* (resp. s_2 *simulates* s_1) and write $s_1 \sim s_2$ (resp. $s_1 \precsim s_2$). ▲

We lift bisimulation and simulation relations to two different SMDPs by considering the disjoint union of the two and comparing their initial states. We denote by $\sim$ the largest bisimulation relation and by $\precsim$ the largest simulation relation. Furthermore, we say that U and V are *equally fast* and write $U \equiv V$ if $U \preceq V$ and $V \preceq U$.

Example 4.9. Consider the two SMDPs U and V in Figure 3 with the same probability measure μ in all states. It is easy to see that U is bisimilar to V, and hence V also simulates U. However, we show that $U \not\preceq V$ in the following way. Construct the scheduler σ for V by letting

$$\sigma(v_0)(a) = 0.5, \ \sigma(v_0)(b) = 0.5, \ \sigma(v_1)(a) = 1, \ \text{and} \ \sigma(v_2)(b) = 1.$$

Now, for any scheduler σ' for U, we must have either $\sigma'(u_0)(a) < 1$ or $\sigma'(u_0)(b) < 1$. If $\sigma'(u_0)(a) < 1$, then

$$\sigma'(u_0)(a) > (\sigma'(u_0)(a))^2 > \cdots > (\sigma'(u_0)(a))^n.$$

Furthermore, we see that

$$\mathbb{P}^\sigma(v_0)(\mathfrak{C}(a^n, t)) = 0.5 \cdot \mu^{*n}(t) \ \text{and} \ \mathbb{P}^{\sigma'}(u_0)(\mathfrak{C}(a^n, t)) = (\sigma'(u_0)(a))^n \cdot \mu^{*n}(t)$$

for $n > 1$. Take some n such that $(\sigma'(u_0)(a))^n < 0.5$ to obtain $\mathbb{P}^{\sigma'}(u_0)(\mathfrak{C}(a^n, t)) < \mathbb{P}^\sigma(v_0)(\mathfrak{C}(a^n, t))$. The same procedure can be used in case $\sigma'(u_0)(b) < 1$. Hence we conclude that $U \not\preceq V$, and therefore also that $U \not\equiv V$. ♦

Example 4.9 also works for schedulers with memory, although the argument has to be modified a bit. In that case, in each step either the probability of a trace consisting only of *a*'s or the probability of a trace consisting only of *b*'s must decrease in U, so after some number of steps, the probability of one of these two must decrease below 0.5, and then the rest of the argument is the same.

Example 4.10. Consider the SMDPs U and V in Figure 2 and let $F_\mu = Exp[\theta_1]$ and $F_\nu = Exp[\theta_2]$ be exponential distributions with rates $\theta_1 > \theta_2 > 0$. Then, as shown in Example 4.7, it holds that $U \preceq V$. However, we have both $U \not\precsim V$ and $U \not\sim V$.   ◆

From Examples 4.9 and 4.10, we get the following theorem.

Theorem 4.11. $\precsim$ *and* $\preceq$ *are incomparable,* $\sim$ *and* $\preceq$ *are incomparable, and we have* $\sim \not\subseteq \equiv$.

Note that Theorem 4.11 holds both for memoryless and for memoryful schedulers that can remember actions and states, but it is still unclear what happens if schedulers are allowed to remember time.

5   Compositionality

Next we introduce the notion of composition of SMDPs. As argued in [25], the style of synchronous CSP composition is the most natural one to consider for reactive probabilistic systems, so this is the one we will adopt. However, we leave the composition of the residence-times as a parameter, so that we can compare different kinds of composition.

Definition 5.1. A function $\star : \mathscr{D}(\mathbb{R}_{\geq 0}) \times \mathscr{D}(\mathbb{R}_{\geq 0}) \to \mathscr{D}(\mathbb{R}_{\geq 0})$ is called a *residence-time composition function* if it is commutative, i.e. $\star(\mu, \nu) = \star(\nu, \mu)$ for all $\mu, \nu \in \mathscr{D}(\mathbb{R}_{\geq 0})$.   ▲

One example of such a composition function is when $\star$ is a coupling, which is a joint probability measure such that its marginals are μ and ν. A simple special case of this is the product measure $\star(\mu, \nu) = \mu \times \nu$, which is defined by $(\mu \times \nu)(B_1 \times B_2) = \mu(B_1) \cdot \nu(B_2)$ for all Borel B_1 and B_2.

In order to model the situation in which we want the composite system only to take a transition when both components can take a transition, it is natural to take the minimum of the two probabilities, which corresponds to waiting for the slowest of the two. In that case, we let $F_{\star(\mu,\nu)}(t) = \min\{F_\mu(t), F_\nu(t)\}$, and we call this *minimum composition*. Likewise, if we only require one of the components to be able to take a transition, then it is natural to take the maximum of the two probabilities by letting $F_{\star(\mu,\nu)}(t) = \max\{F_\mu(t), F_\nu(t)\}$, which we call *maximum composition*. A special case of minimum composition is the composition on rates used in PEPA [11], and a special case of maximum composition is the composition on rates used in TIPP [8].

Further knowledge about the processes that are being composed lets one define more specific composition functions. As an example, if we know that the components only have exponential distributions, then we can define composition functions that work directly on the rates of the distributions. If $F_\mu = Exp[\theta]$ and $F_\nu = Exp[\theta']$, then one could for example let $\star(\mu, \nu)$ be such that $F_{\star(\mu,\nu)} = Exp[\theta \cdot \theta']$. This corresponds to the composition on rates that is used in SPA [10], and we will call it *product composition*. Note that product composition is *not* given by the product measure.

Definition 5.2. Let $\star$ be a residence-time composition function. Then the $\star$-*composition* of U and V, denoted by $U \parallel_\star V = (S, s_0, \tau, \rho)$, is given by

- $S = U \times V$,

- $s_0 = (u_0, v_0)$,

- $\tau((u,v), a)((u', v')) = \tau_U(u, a)(u') \cdot \tau_V(v, a)(v')$ for all $a \in L$ and $(u', v') \in S$, and

- $\rho((u, v)) = \star(\rho_U(u), \rho_V(v))$.   ▲

We write $u \parallel_\star v$ to denote the composite state (u, v) of $U \parallel_\star V$ where $u \in S_U$ and $v \in S_V$.
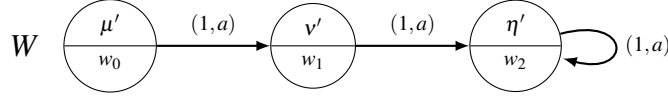
Figure 4: For different instantiations of μ', ν', and η', the context W leads to parallel timing anomalies for product, minimum, and maximum rate composition, respectively.

5.1 Parallel Timing Anomalies

If we have two components U and V, and we know that U is faster than V, then if V is in parallel with some context W, we would expect this composition to become faster when we replace the component V with the component U. However, sometimes this fails to happen, and we will call such an occurrence a *parallel timing anomaly*.

In this section we show that parallel timing anomalies can occur for some of the kinds of composition discussed in Section 5. We do this by giving different contexts W for the SMDPs U and V from Figure 2, for which it was shown in Example 4.7 that $U \preceq V$. Our examples of parallel timing anomalies make no use of non-determinism or probabilistic branching, thus showing that the parallel timing anomalies are caused inherently by the timing behaviour of the SMDPs. For ease of presentation, we let the set of labels L consist only of the label a in this section.

Consider the two SMDPs U and V depicted in Figure 2 and let the context W be given by Figure 4. For the examples in this section, let $F_\mu = Exp[2]$, $F_\nu = Exp[0.5]$, and let $\eta = \eta'$ be arbitrary. From Example 4.7 we know that $U \preceq V$.

Example 5.3 (Product composition). Let $\star$ be product composition and let $F_{\mu'} = Exp[10]$ and $F_{\nu'} = Exp[0.1]$. In $U \parallel_\star W$, the rates in the first two states will then be 20 and 0.05, and in $V \parallel_\star W$ they will be 5 and 0.2. Consider the time-bounded cylinder $\mathfrak{C}(aa,2)$. Then we see that

$$\mathbb{P}(u_0 \parallel_\star w_0)(\mathfrak{C}(aa,2)) \approx 0.09 \quad \text{and} \quad \mathbb{P}(v_0 \parallel_\star w_0)(\mathfrak{C}(aa,2)) \approx 0.30,$$

showing that $U \parallel_\star W \npreceq V \parallel_\star W$. Hence we have a parallel timing anomaly. What happens is that in the process $V \parallel_\star W$ the probability of taking a transition before time 2 with rate 5 is already very close to 1, so the process $U \parallel_\star W$ does not gain much by having a rate of 20, whereas in the next step, $V \parallel_\star W$ gains a lot of probability by having a rate of 0.5 compared to the rate 0.05 of $U \parallel_\star W$. ◆

Example 5.4 (Minimum composition). Let $\star$ be minimum composition and let $F_{\mu'} = Exp[1]$ and $F_{\nu'} = Exp[2]$. The rates of $U \parallel_\star W$ are then 1 and 0.5, whereas they are 0.5 and 2 in $V \parallel_\star W$. Then we get

$$\mathbb{P}(u_0 \parallel_\star w_0)(\mathfrak{C}(aa,2)) \approx 0.40 \quad \text{and} \quad \mathbb{P}(v_0 \parallel_\star w_0)(\mathfrak{C}(aa,2)) \approx 0.51,$$

so $U \parallel_\star W \npreceq V \parallel_\star W$. What happens in this example is that in the second step, $U \parallel_\star W$ has the same rate as $V \parallel_\star W$ had in the first step. This means that $U \parallel_\star W$ must be proportionally faster in the second step. However, $V \parallel_\star W$ has a rate of 2 in the second step, but $U \parallel_\star W$ only had a rate of 1 in the first step. ◆

Example 5.5 (Maximum composition). Let $\star$ be maximum composition and let $F_{\mu'} = Exp[2]$ and $F_{\nu'} = Exp[1]$. $U \parallel_\star W$ then has rates 2 and 1, and $V \parallel_\star W$ has rates 2 and 2. Then

$$\mathbb{P}(u_0 \parallel_\star w_0)(\mathfrak{C}(aa,2)) \approx 0.75 \quad \text{and} \quad \mathbb{P}(v_0 \parallel_\star w_0)(\mathfrak{C}(aa,2)) \approx 0.91,$$

so $U \parallel_\star W_3 \npreceq V \parallel_\star W_3$. The reason for the timing anomaly in this case is clear: $V \parallel_\star W$ simply has a higher rate in each step than $U \parallel_\star W$ does. ◆

5.2   Avoiding Parallel Timing Anomalies

We have seen in the previous section that parallel timing anomalies can occur. We now wish to understand what kind of contexts do not lead to timing anomalies. In this section we assume that the set L of transition labels is a finite set. Also, we fix a residence-time composition function $\star$ and two additional SMDPs $W = (S_W, w_0, \tau_W, \rho_W)$ and $W' = (S_{W'}, w'_0, \tau_{W'}, \rho_{W'})$ which should be thought of as contexts. Next we identify conditions on (W, w_0) such that $U \preceq V$ will imply $U \parallel_\star W \preceq V \parallel_\star W$.

   We first give conditions that over-approximate the faster-than relation between the composite systems by requiring that when U and W are put in parallel, then the composite system is point-wise faster than U along all paths. Likewise, we require that when V and W are put in parallel, the composite system is point-wise slower than V along all paths. If we already know that U is faster than V, this will imply by transitivity that $U \parallel_\star W$ is faster than $V \parallel_\star W$. We have already seen in Example 4.7 that a process U need not be point-wise faster than V along all paths in order for U to be faster than V. However, by imposing this condition, we do not need to compare convolutions of distributions, but can compare the distributions directly.

   We will say that an SMDP M has a *deterministic Markov kernel* if for all states s and labels a, there is at most one state s' such that $\tau(s, a)(s') > 0$. A *state path* of length n in M is a sequence $s_1 \ldots s_n$ such that s_1 is the initial state and for all i there exists $a \in L$ such that $\tau(s_i, a)(s_{i+1}) > 0$.

Definition 5.6. Let $n \in \mathbb{N}$. We say that $\star$ is *n-monotonic in U, V, W, and W'*, written $(U, W) \lesssim^n_\star (V, W')$, if W' has a deterministic Markov kernel and for all state paths $u_1 \ldots u_n$, $v_1 \ldots v_n$, $w_1 \ldots w_n$, and $w'_1 \ldots w'_n$ we have

- $F_{\rho(u_i \parallel_\star w_i)}(t) \geq F_{\rho_U(u_i)}(t)$ and $F_{\rho_V(v_i)}(t) \geq F_{\rho(v_i \parallel_\star w'_i)}(t)$ for all $t \in \mathbb{R}_{\geq 0}$ and $1 \leq i \leq n$,

- for all schedulers σ_U for U there exists a scheduler $\sigma_{U,W}$ for $U \parallel_\star W$ such that we have

$$\tau^{\sigma_{U,W}}(u_i \parallel_\star w_i, a)(u_{i+1} \parallel_\star w_{i+1}) \geq \tau^{\sigma_U}_U(u_i, a)(u_{i+1}), \quad \text{and}$$

- for all schedulers $\sigma_{V,W'}$ for $V \parallel_\star W'$ there exists a scheduler σ_V for V such that we have

$$\tau^{\sigma_V}_V(v_i, a)(v_{i+1}) \geq \tau^{\sigma_{V,W'}}(v_i \parallel_\star w'_i, a)(v_{i+1} \parallel_\star w'_{i+1})$$

and for all $a \in L$ and $1 \leq i < n$. Furthermore, we will say that $\star$ is *monotonic in U, V, W, and W'* and write $(U, W) \lesssim_\star (V, W')$, if it is n-monotonic in $U, V, W,$ and W' for all $n \in \mathbb{N}$.   ▲

   Clearly, if $(U, W) \lesssim^n_\star (V, W')$, then $(U, W) \lesssim^m_\star (V, W')$ for all $m \leq n$. The next result shows that if $(U, W) \lesssim_\star (V, W')$, then we are guaranteed to avoid parallel timing anomalies.

Theorem 5.7. *If $(U, W) \lesssim_\star (V, W')$ as well as $U \preceq V$ and $W \preceq W'$, then we have $U \parallel_\star W \preceq V \parallel_\star W'$.*

   The special case where $W = W'$ shows that this condition is sufficient to avoid parallel timing anomalies. We do not know if it is decidable whether $(U, W) \lesssim_\star (V, W')$. However, there is a stronger condition which is decidable in the case of finite SMDPs. We present it in the next definition.

Definition 5.8. We say that $\star$ is *strongly n-monotonic in U, V, W, and W'* and write $(U, W) \leq^n_\star (V, W')$ if W' has a deterministic Markov kernel and for all state paths $u_1 \ldots u_n$, $v_1 \ldots v_n$, $w_1 \ldots w_n$, and $w'_1 \ldots w'_n$,

- $F_{\rho(u_i \parallel_\star w_i)}(t) \geq F_{\rho_U(u_i)}(t)$ and $F_{\rho_V(v_i)}(t) \geq F_{\rho(v_i \parallel_\star w'_i)}(t)$ for all $t \in \mathbb{R}_{\geq 0}$ and $1 \leq i \leq n$

- for all schedulers σ_U for U and all schedulers $\sigma_{U,W}$ for $U \parallel_\star W$ we have

$$\tau^{\sigma_{U,W}}(u_i \parallel_\star w_i, a)(u_{i+1} \parallel_\star w_{i+1}) \geq \tau^{\sigma_U}_U(u_i, a)(u_{i+1}), \quad \text{and}$$

- for all schedulers $\sigma_{V,W'}$ for $V \parallel_\star W'$ and all schedulers σ_V for V we have

$$\tau_V^{\sigma_V}(v_i, a)(v_{i+1}) \geq \tau^{\sigma_{V,W'}}(v_i \parallel_\star w_i', a)(v_{i+1} \parallel_\star w_{i+1}')$$

for all $a \in L$ and $1 \leq i < n$. If $(U,W) \leq_\star^n (V,W')$ for all $n \in \mathbb{N}$, we say that $\star$ is *strongly monotonic* in U, V, W, and W' and write $(U,W) \leq_\star (V,W')$.   ▲

The conditions of Definition 5.8 are similar to the conditions from Definition 5.6, but the existential quantifier is strengthened to a universal quantifier. It is obvious that $(U,W) \leq_\star (V,W')$ implies $(U,W) \lesssim_\star (V,W')$, and hence we get the following corollary.

Corollary 5.9. *If $(U,W) \leq_\star (V,W')$ as well as $U \preceq V$ and $W \preceq W'$, then $U \parallel_\star W \preceq V \parallel_\star W'$.*

Example 5.10. Let U and V be given by Figure 2 with $F_\mu \geq F_\nu$ as in Example 4.7. Let $\star$ be minimum rate composition and consider the context W from Figure 4, where $\mu' = \mu$, $\nu' = \nu$, and $\eta' = \eta$. There is only one possible scheduler σ, which is the Dirac measure at a, and hence it is clear that the second and third conditions are satisfied. We also find that

$$F_{\rho(u_0 \parallel_\star w_0)}(t) = F_{\rho_U(u_0)}(t) \qquad\qquad F_{\rho_V(v_0)}(t) = F_{\rho(v_0 \parallel_\star w_0)}(t)$$
$$F_{\rho(u_1 \parallel_\star w_1)}(t) = F_{\rho_U(u_1)}(t) \qquad\qquad F_{\rho_V(v_1)}(t) = F_{\rho(v_1 \parallel_\star w_1)}(t)$$
$$F_{\rho(u_2 \parallel_\star w_2)}(t) = F_{\rho_U(u_2)}(t) \qquad\qquad F_{\rho_V(v_2)}(t) = F_{\rho(v_2 \parallel_\star w_2)}(t)$$

and hence the first condition is also satisfied, so $(U,W) \leq_\star (V,W)$.   ◆

Example 5.11. All the examples we gave in Section 5.1 fail to be monotonic, and hence also fail to be strongly monotonic, since they all violate the first condition of monotonicity.

In Example 5.3, this is because

$$F_{\rho(v_0)}(t) = Exp[0.5](t) < Exp[5](t) = F_{\rho(v_0 \parallel_\star w_0)}(t) \quad \text{for any } t > 0.$$

Likewise, in Example 5.4 we have

$$F_{\rho(u_0 \parallel_\star w_0)}(t) = Exp[1](t) < Exp[2](t) = F_{\rho(u_0)}(t) \quad \text{for any } t > 0.$$

Finally, in Example 5.5 we have

$$F_{\rho(v_0)}(t) = Exp[0.5](t) < Exp[2](t) = F_{\rho(v_0 \parallel_\star w_0)}(t) \quad \text{for any } t > 0.$$ ◆

The following is the main theorem of our paper, showing that strong monotonicity implies the absence of parallel timing anomalies.

Theorem 5.12. *If $(U,W) \leq_\star (V,W')$ as well as $U \preceq V$ and $W \preceq W'$, then $U \parallel_\star W \preceq V \parallel_\star W'$.*

We now wish to show that it is decidable whether $(U,W) \leq_\star (V,W')$ for finite SMDPs, thereby giving a decidable condition for avoiding timing anomalies. To do this, we first show that in order to establish strong monotonicity, it is enough to consider paths up to length

$$m = \max\{|S_U| \cdot |S_W|, |S_V| \cdot |S_{W'}|\} + \max\{|S_U|, |S_V|, |S_W|, |S_{W'}|\} + 1,$$

due to the fact that they start repeating, as can be shown by a simple pigeonhole argument.

Lemma 5.13. *Let U, V, W, and W' be finite. If $(U,W) \leq_\star^m (V,W')$, then $(U,W) \leq_\star (V,W')$.*

We can now use the first-order theory of the reals to show that strong monotonicity is a decidable property. In order for this to work, we must be able to check the first condition of strong monotonicity in the first-order theory of the reals. We therefore need the sets of $t \in \mathbb{R}_{\geq 0}$ in that condition to be expressible in the first-order theory of reals, which can be accomplished by assuming that these sets are semialgebraic.

Theorem 5.14. *Let U, V, W, and W' be finite. If for all state paths $u_1 \ldots u_n$, $v_1 \ldots v_n$, $w_1 \ldots w_n$, and $w'_1 \ldots w'_n$ we have that $\{t \in \mathbb{R}_{\geq 0} \mid F_{\rho(u_i \|_\star w_i)}(t) \geq F_{\rho_U(u_i)}(t)\}$ and $\{t \in \mathbb{R}_{\geq 0} \mid F_{\rho_V(v_i)}(t) \geq F_{\rho(v_i \|_\star w'_i)}(t)\}$ are semialgebraic sets for all $1 \leq i \leq m$, then it is decidable whether $(U,W) \leq_\star (V,W')$.*

For uniform and exponential distributions with minimum or maximum composition, the corresponding sets are all semialgebraic, and the same is true for exponential distributions with product composition. Theorem 5.14 can therefore be used for these types of composition.

It turns out that strong monotonicity implies the absence of non-determinism.

Proposition 5.15. *If $(U,W) \leq_\star (V,W')$, then L is a singleton set or u_0 is a deadlock state, i.e. the transition probability is zero from u_0 to any other state.*

However, strong monotonicity still makes sense as a condition, since all our examples of timing anomalies in Section 5.1 have no non-determinism.

6   Conclusion

In this paper, we have investigated the notion of a process being faster than another process in the context of semi-Markov decision processes. We have given a trace-based definition of a faster-than relation, and shown how this notion relates to the usual notions of simulation and bisimulation and to convolutions of distributions. By considering composition as being parametric in how the residence times of states are combined, we have given examples showing that our faster-than relation gives rise to parallel timing anomalies for many of the popular ways of composing rates. We have therefore given sufficient conditions for how such parallel timing anomalies can be avoided, and we have shown that these conditions are decidable.

While the conditions for strong monotonicity are so strict that they rule out non-determinism, it seems difficult to obtain better results that are still decidable. This is because the faster-than relation itself is undecidable whenever non-determinism is introduced, and even without non-determinism, it has close connections to an important problem in number theory [21] whose decidability status is a long-standing open problem. Because of this, an interesting future work direction is to explore the boundaries of decidability for conditions to avoid timing anomalies.

Furthermore, the conditions we have given for avoiding timing anomalies do not look at the context in isolation, but depend also on the processes that are being swapped. It would be preferable to have conditions on a context that would guarantee the absence of parallel timing anomalies no matter what processes are being swapped.

References

[1] S. Akshay, Timos Antonopoulos, Joël Ouaknine & James Worrell (2015): *Reachability problems for Markov chains*. Inf. Process. Lett. 115(2), pp. 155–158, doi:10.1016/j.ipl.2014.08.013.

[2] Christel Baier, Joost-Pieter Katoen, Holger Hermanns & Verena Wolf (2005): *Comparative branching-time semantics for Markov chains*. Inf. Comput. 200(2), pp. 149–214, doi:10.1016/j.ic.2005.03.001.

[3] Patrick Billingsley (1995): *Probability And Measure*, 3rd edition. Wiley-Interscience.

[4] Franck Cassez, René Rydhof Hansen & Mads Chr. Olesen (2012): *What is a Timing Anomaly?* In Tullio Vardanega, editor: *12th International Workshop on Worst-Case Execution Time Analysis, WCET 2012, July 10, 2012, Pisa, Italy*, OASICS 23, Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, pp. 1–12, doi:10.4230/OASIcs.WCET.2012.1.

[5] Edmund M. Clarke, David E. Long & Kenneth L. McMillan (1989): *Compositional Model Checking*. In: *Proceedings of the Fourth Annual Symposium on Logic in Computer Science (LICS '89), Pacific Grove, California, USA, June 5-8, 1989*, IEEE Computer Society, pp. 353–362, doi:10.1109/LICS.1989.39190.

[6] Flavio Corradini, Roberto Gorrieri & Marco Roccetti (1995): *Performance Preorder: Ordering Processes with Respect to Speed*. In Jirí Wiedermann & Petr Hájek, editors: *Mathematical Foundations of Computer Science 1995, 20th International Symposium, MFCS'95, Prague, Czech Republic, August 28 - September 1, 1995, Proceedings, Lecture Notes in Computer Science* 969, Springer, pp. 444–453, doi:10.1007/3-540-60246-1_150.

[7] Marc Geilen, Stavros Tripakis & Maarten Wiggers (2011): *The earlier the better: a theory of timed actor interfaces*. In Marco Caccamo, Emilio Frazzoli & Radu Grosu, editors: *Proceedings of the 14th ACM International Conference on Hybrid Systems: Computation and Control, HSCC 2011, Chicago, IL, USA, April 12-14, 2011*, ACM, pp. 23–32, doi:10.1145/1967701.1967707.

[8] Norbert Götz, Ulrich Herzog & Michael Rettelbach (1993): *Multiprocessor and Distributed System Design: The Integration of Functional Specification and Performance Analysis Using Stochastic Process Algebras*. In Lorenzo Donatiello & Randolph D. Nelson, editors: *Performance Evaluation of Computer and Communication Systems, Joint Tutorial Papers of Performance '93 and Sigmetrics '93, Santa Clara, CA, USA, May 10-14, 1993, Lecture Notes in Computer Science* 729, Springer, pp. 121–146, doi:10.1007/BFb0013851.

[9] Ronald L. Graham (1969): *Bounds on Multiprocessing Timing Anomalies*. SIAM Journal of Applied Mathematics 17(2), pp. 416–429, doi:10.1137/0117039.

[10] Holger Hermanns, Ulrich Herzog & Vassilis Mertsiotakis (1998): *Stochastic Process Algebras - Between LOTOS and Markov Chains*. Computer Networks 30(9-10), pp. 901–924, doi:10.1016/S0169-7552(97)00133-5.

[11] Jane Hillston (1996): *A compositional approach to performance modelling*. Distinguished Dissertations in Computer Science, Cambridge University Press, doi:10.1017/CBO9780511569951.

[12] Raimund Kirner, Albrecht Kadlec & Peter P. Puschner (2009): *Precise Worst-Case Execution Time Analysis for Processors with Timing Anomalies*. In: *21st Euromicro Conference on Real-Time Systems, ECRTS 2009, Dublin, Ireland, July 1-3, 2009*, IEEE Computer Society, pp. 119–128, doi:10.1109/ECRTS.2009.8.

[13] Edward A. Lee (2008): *Cyber Physical Systems: Design Challenges*. In: *11th IEEE International Symposium on Object-Oriented Real-Time Distributed Computing (ISORC 2008), 5-7 May 2008, Orlando, Florida, USA*, IEEE Computer Society, pp. 363–369, doi:10.1109/ISORC.2008.25.

[14] Thomas Lundqvist & Per Stenström (1999): *Timing Anomalies in Dynamically Scheduled Microprocessors*. In: *Proceedings of the 20th IEEE Real-Time Systems Symposium, Phoenix, AZ, USA, December 1-3, 1999*, IEEE Computer Society, pp. 12–21, doi:10.1109/REAL.1999.818824.

[15] Gerald Lüttgen & Walter Vogler (2001): *A Faster-than Relation for Asynchronous Processes*. In Kim Guldstrand Larsen & Mogens Nielsen, editors: *CONCUR 2001 - Concurrency Theory, 12th International Conference, Aalborg, Denmark, August 20-25, 2001, Proceedings, Lecture Notes in Computer Science* 2154, Springer, pp. 262–276, doi:10.1007/3-540-44685-0_18.

[16] Gerald Lüttgen & Walter Vogler (2006): *Bisimulation on speed: A unified approach*. Theor. Comput. Sci. 360(1-3), pp. 209–227, doi:10.1016/j.tcs.2006.03.004.

[17] Sneana Maovi, Saa Stoi & Rade Hajdin (2015): *Application of Semi-Markov Decision Process in Bridge Management*. IABSE Symposium Report 105(28), pp. 1–8, doi:10.2749/222137815818358204.

[18] Faron Moller & Chris M. N. Tofts (1991): *Relating Processes With Respect to Speed*. In Jos C. M. Baeten & Jan Friso Groote, editors: *CONCUR '91, 2nd International Conference on Concurrency Theory, Amsterdam, The Netherlands, August 26-29, 1991, Proceedings*, Lecture Notes in Computer Science 527, Springer, pp. 424–438, doi:10.1007/3-540-54430-5_104.

[19] Martin R. Neuhäußer & Joost-Pieter Katoen (2007): *Bisimulation and Logical Preservation for Continuous-Time Markov Decision Processes*. In Luís Caires & Vasco Thudichum Vasconcelos, editors: *CONCUR 2007 - Concurrency Theory, 18th International Conference, CONCUR 2007, Lisbon, Portugal, September 3-8, 2007, Proceedings*, Lecture Notes in Computer Science 4703, Springer, pp. 412–427, doi:10.1007/978-3-540-74407-8_28.

[20] Joël Ouaknine & James Worrell (2014): *Positivity Problems for Low-Order Linear Recurrence Sequences*. In Chandra Chekuri, editor: *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2014, Portland, Oregon, USA, January 5-7, 2014*, SIAM, pp. 366–379, doi:10.1137/1.9781611973402.27.

[21] Mathias Ruggaard Pedersen, Nathanaël Fijalkow, Giorgio Bacci, Kim G. Larsen & Radu Mardare (2018): *Timed Comparisons of Semi-Markov Processes*. In Shmuel Tomi Klein, Carlos Martín-Vide & Dana Shapira, editors: *Language and Automata Theory and Applications - 12th International Conference, LATA 2018, Ramat Gan, Israel, April 9-11, 2018, Proceedings*, Lecture Notes in Computer Science 10792, Springer, pp. 271–283, doi:10.1007/978-3-319-77313-1_21.

[22] Muhammad Reza Pulungan (2009): *Reduction of Acyclic Phase-Type Representations*. Ph.D. thesis, Faculty of Natural Sciences and Technology, Saarland University, doi:10.22028/D291-25951.

[23] Jan Reineke, Björn Wachter, Stephan Thesing, Reinhard Wilhelm, Ilia Polian, Jochen Eisinger & Bernd Becker (2006): *A Definition and Classification of Timing Anomalies*. In: *WCET*, OASICS 4, Internationales Begegnungs- und Forschungszentrum fuer Informatik (IBFI), Schloss Dagstuhl, Germany, doi:10.4230/OASIcs.WCET.2006.671.

[24] Ichiro Satoh & Mario Tokoro (1994): *A Formalism for Remotely Interacting Processes*. In Takayasu Ito & Akinori Yonezawa, editors: *Theory and Practice of Parallel Programming, International Workshop TPPP'94, Sendai, Japan, November 7-9, 1994, Proceedings*, Lecture Notes in Computer Science 907, Springer, pp. 216–228, doi:10.1007/BFb0026571.

[25] Ana Sokolova & Erik P. de Vink (2004): *Probabilistic Automata: System Types, Parallel Composition and Comparison*. In Christel Baier, Boudewijn R. Haverkort, Holger Hermanns, Joost-Pieter Katoen & Markus Siegle, editors: *Validation of Stochastic Systems - A Guide to Current Research*, Lecture Notes in Computer Science 2925, Springer, pp. 1–43, doi:10.1007/978-3-540-24611-4_1.

[26] Terence Tao (2013): *An Introduction to Measure Theory*. Graduate studies in mathematics, American Mathematical Society.

[27] O. S. Thomas & J. O. Sobanjo (2016): *Semi-Markov Decision Process: A Decision Tool for Transportation Infrastructure Management Systems*. In: *International Conference on Transportation and Development*, pp. 384–396, doi:10.1061/9780784479926.036.

[28] Walter Vogler (1995): *Timed Testing of Concurrent Systems*. Inf. Comput. 121(2), pp. 149–171, doi:10.1006/inco.1995.1130.

[29] Walter Vogler (2003): *Faster asynchronous systems*. Inf. Comput. 184(2), pp. 311–342, doi:10.1016/S0890-5401(03)00065-8.