Compositionality of Approximate Bisimulation for Probabilistic Systems

Daniel Gebler
Department of Computer Science, VU University Amsterdam,
De Boelelaan 1081a, NL-1081 HV Amsterdam, The Netherlands
e.d.gebler@vu.nl

Simone Tini
Department of Scienza e Alta Tecnologia,
University of Insubria, Via Valleggio 11, I-22100, Como, Italy
simone.tini@uninsubria.it

Probabilistic transition system specifications using the rule format $nt\mu f\theta/nt\mu x\theta$ provide structural operational semantics for Segala-type systems and guarantee that probabilistic bisimilarity is a congruence. Probabilistic bisimilarity is for many applications too sensitive to the exact probabilities of transitions. Approximate bisimulation provides a robust semantics that is stable with respect to implementation and measurement errors of probabilistic behavior. We provide a general method to quantify how much a process combinator expands the approximate bisimulation distance. As a direct application we derive an appropriate rule format that guarantees compositionality with respect to approximate bisimilarity. Moreover, we describe how specification formats for non-standard compositionality requirements may be derived.

1 Introduction

Over the last decade a number of researchers have started to develop a theory of structural operational semantics (SOS) [17] for probabilistic transition systems (PTSs). Several rule formats for various PTSs were proposed that ensure compositionality (in technical terms congruence) of probabilistic bisimilarity [1,2,4,12,13,15]. The rule format $nt\mu f\theta/nt\mu x\theta$ [4,15] subsumes all earlier formats and can be understood as the probabilistic variant of the *ntyft/ntyxt* format [10].

Probabilistic bisimilarity is very sensitive to the exact probabilities of transitions. The slightest perturbation of the probabilities can destroy bisimilarity. Two proposals for a more robust semantics of probabilistic processes have been put forward. The *metric bisimulation* approach [3, 6, 8] is the quantitative analogue of the relational notion of probabilistic bisimulation. It assigns a distance to each pair of processes, which measures the proximity of their quantitative properties. Another approach is the *approximate bisimulation* (also called ε-bisimulation) approach [7, 8, 22]. Approximate bisimulations are probabilistic bisimulations where the transfer condition is relaxed, namely two processes are related by an ε-bisimulation if their probability to reach a set of states related by that ε-bisimulation differs by at most ε. Processes that are related by an ε-bisimulation with ε being small are "almost bisimiliar". Approximate bisimulations are not transitive in general, as two states with quite different behaviors could be linked by a sequence of states, which pairwise have only little behavioral difference. Approximate bisimulations have been characterized in operational terms [8], by a modal logic [7, 22], and in terms of games [7]. The metric and approximate bisimulation approach are in general not comparable (see [20–22]). The main difference is that in the approximate bisimulation approach (contrary to the

J. Borgström and B. Luttik (Eds.): Combined Workshop on Expressiveness in Concurrency and Structural Operational Semantics (EXPRESS/SOS 2013)
EPTCS 120, 2013, pp. 32–46, doi:10.4204/EPTCS.120.4

metric bisimulation approach) the differences along paths are neither accumulated nor weighted by the probability of the realization of that path. In this paper we consider approximate bisimulations.

In order to allow for compositional specification and reasoning, it is necessary that the considered behavioral semantics is compatible with all operators of the language of interest. For behavioral equivalences (e.g. probabilistic bisimulations) this is the well-known congruence property. For approximate bisimulations the quantitative analogue to the congruence property requires that when different processes are combined by a process combinator (i.e., an operator of the language), then the distance between the resulting combined processes is (reasonably) bounded. A natural notion for this bound is the sum of the distances between the processes to be combined [6]. A process combinator respecting this specific bound is called *non-expansive*. Intuitively, this bound expresses that a process combinator does not increase the behavioral distance of the processes to be combined. The congruence property and non-expansivity property of an n-ary process combinator f can be expressed by the following proof rules (with $\sim$ denoting the probabilistic bisimilarity and d denoting the approximate bisimulation distance):

$$\frac{s_i \sim t_i \text{ for all } i = 1,\ldots,n}{f(s_1,\ldots,s_n) \sim f(t_1,\ldots,t_n)} \qquad \frac{d(s_i,t_i) \leq \varepsilon_i \text{ for all } i = 1,\ldots,n}{d(f(s_1,\ldots,s_n),f(t_1,\ldots,t_n)) \leq \sum_{i=1}^{n} \varepsilon_i}$$

However, for specific applications, alternative compositionality requirements are required that allow for more or less variance (than the linear sum used in non-expansivity) of the combined processes. For instance, a process combinator that combines a number of distributed systems with a measurement unit may allow for some variance in the combined distributed systems, but must enforce that the measurement unit itself is strict.

In this paper we report a substantial first step towards a theory of robust specifications for probabilistic processes. As an operational model for probabilistic processes, we consider Segala-type PTSs that exhibit both probabilistic and nondeterministic behavior. The probabilistic processes are specified by probabilistic transition system specifications (PTSS) with simple $nt\mu f\theta/nt\mu x\theta$ rules. By simple $nt\mu f\theta/nt\mu x\theta$ rules we mean $nt\mu f\theta/nt\mu x\theta$ rules without lookahead. In order to facilitate compositional specification and reasoning, we study how the distance between two terms with the same topmost function symbol depends on the distances of the arguments. In detail, we characterize the *expansivity* of a process combinator, which gives an upper bound on the distance of the combined processes given the distance between their components. Formally, the expansivity of a process combinator f with n arguments is defined as a mapping $\mathbb{R}^n \to \mathbb{R}$ taking distances of the arguments $\varepsilon_1,\ldots\varepsilon_n$ to ε, with ε defined as the maximal distance between all $f(s_1,..,s_n)$ and $f(t_1,..,t_n)$ whenever all s_i and t_i are in approximate bisimulation distance ε_i.

The first contribution of our paper is the characterization of the expansivity of each process combinator. The expansivity of a process combinator is defined as the least fixed point of a monotone function that counts recursively how often the processes are copied. Our second contribution is to deduce, from the expansivity of process combinators, an appropriate rule format that guarantees non-expansivity of all operators specified in this format. The rule format is derived from the simple $nt\mu f\theta/nt\mu x\theta$ rule format by prohibiting that source processes or derivatives are copied. Finally, we demonstrate how the expansivity of process combinators can be used to derive rule formats for alternative compositionality requirements.

We consider in this paper approximate bisimulations because the relaxed transfer condition preserves the basic relational nature of probabilistic bisimulations and allows us to apply (adapted and extended) known proof techniques developed for congruence rule formats. Moreover, the new techniques introduced in this paper to quantify the expansivity of process combinators translate naturally to bisimulation

metrics. In this sense, we are also opening the door to develop a theory of robust process specifications with respect to bisimulation metrics.

This is the first paper that explores systematically the approximate bisimulation distance of probabilistic processes specified by transition system specifications. Tini already proposed a rule format for reactive probabilistic processes [19, 20]. Our format significantly generalizes and extends that format. First of all, we apply the more general Segala-type systems that admit, besides probabilistic behavior (probabilistic choice), nondeterministic reactive behavior (internal nondeterministic branching). Furthermore, while Tini used a notion of approximate bisimulation, which is an equivalence but not closed under union, we are using the (by now) standard notion [7, 22], which is only reflexive and symmetric but closed under union. Finally, the novel rule format based on counting of copies of processes and their derivatives in its defining rules allows us to handle a wider class of process combinators that ensure non-expansivity.

2 Preliminaries

We assume an infinite set of (state) variables $\mathscr{V}$. We let x, y, z range over $\mathscr{V}$. A *signature* is a structure $\Sigma = (F, \mathsf{r})$, where (i) F is a set of *function names* (operators) disjoint from $\mathscr{V}$, and (ii) $\mathsf{r} : F \to \mathbb{N}$ is a *rank function*, which gives the arity of a function name. An operator $f \in F$ is called a *constant* if $\mathsf{r}(f) = 0$. We write $f \in \Sigma$ for $f \in F$. Let $W \subseteq \mathscr{V}$ be a set of variables. The set of Σ-terms (also called state terms) over W, denoted by $T(\Sigma, W)$, is the least set satisfying: (i) $W \subseteq T(\Sigma, W)$, and (ii) if $f \in \Sigma$ and $t_1, \cdots, t_{\mathsf{r}(f)} \in T(\Sigma, W)$, then $f(t_1, \cdots, t_{\mathsf{r}(f)}) \in T(\Sigma, W)$. $T(\Sigma, \emptyset)$ is the set of all *closed terms* and abbreviated as $T(\Sigma)$. $T(\Sigma, \mathscr{V})$ is the set of *open terms* and abbreviated as $\mathbb{T}(\Sigma)$. We may refer to operators as process combinators, and refer to terms as processes. $\mathrm{Var}(t) \subseteq \mathscr{V}$ denotes the set of variables in t. $\mathrm{MVar} \colon \mathbb{T}(\Sigma) \to (\mathscr{V} \to \mathbb{N})$ denotes for $\mathrm{MVar}(t)(x)$ how often the variable x occurs in t. A (state variable) *substitution* is a mapping $\sigma_{\mathscr{V}} : \mathscr{V} \to \mathbb{T}(\Sigma)$. A substitution is closed if it maps each variable to a closed term. A substitution extends to a mapping from terms to terms as usual.

Let $\Delta(T(\Sigma))$ denote the set of all (discrete) probability distributions on $T(\Sigma)$. We let π range over $\Delta(T(\Sigma))$. For $S \subseteq T(\Sigma)$ we define $\pi(S) = \sum_{t \in S} \pi(t)$. For each $t \in T(\Sigma)$, let δ_t denote the *Dirac distribution*, i.e., $\delta_t(t) = 1$ and $\delta_t(t') = 0$ if t and t' are not syntactically equal. The convex combination $\sum_{i \in I} p_i \pi_i$ of a family $\{\pi_i\}_{i \in I}$ of probability distributions with $p_i \in (0, 1]$ and $\sum_{i \in I} p_i = 1$ is defined by $(\sum_{i \in I} p_i \pi_i)(t) = \sum_{i \in I} (p_i \pi_i(t))$. By $f(\pi_1, \ldots, \pi_{\mathsf{r}(f)})$ we denote the distribution that is defined by $f(\pi_1, \ldots, \pi_{\mathsf{r}(f)})(f(t_1, \ldots, t_{\mathsf{r}(f)})) = \prod_{i=1}^{\mathsf{r}(f)} \pi_i(t_i)$. We may use the infix notation where appropriate.

In order to describe probabilistic behavior, we need expressions that denote probability distributions. We assume an infinite set of distribution variables $\mathscr{D}$. We let μ range over $\mathscr{D}$ and ζ range over $\mathscr{D} \cup \mathscr{V}$. Let $D \subseteq \mathscr{D}$ be a set of distribution variables and $V \subseteq \mathscr{V}$ be a set of state variables. The set of *distribution terms* over D and V, notation $\mathsf{DT}(\Sigma, D, V)$, is the least set satisfying: (i) $D \cup \{\delta_t \mid t \in T(\Sigma, V)\} \subseteq \mathsf{DT}(\Sigma, D, V)$, (ii) $\sum_{i \in I} p_i \theta_i \in \mathsf{DT}(\Sigma, D, V)$ if $\theta_i \in \mathsf{DT}(\Sigma, D, V)$ and $p_i \in (0, 1]$ with $\sum_{i \in I} p_i = 1$, and (iii) $f(\theta_1, \ldots, \theta_{\mathsf{r}(f)}) \in \mathsf{DT}(\Sigma, D, V)$ if $f \in \Sigma$ and $\theta_i \in \mathsf{DT}(\Sigma, D, V)$.[1] A *distribution variable* $\mu \in D$ is a variable that takes values from $\Delta(T(\Sigma))$. An *instantiable Dirac distribution* δ_t with $t \in \mathbb{T}(\Sigma)$ is a symbol that takes value $\delta_{t'}$ when variables in t are substituted so that t becomes the closed term $t' \in T(\Sigma)$. Case ii allows one to construct convex combinations of distributions. For concrete terms we use the infix notation, e.g., $[p_1]\theta_1 \oplus [p_2]\theta_2$

[1] This fixes a flaw in [4, 15] where arbitrary functions $f \colon T(\Sigma)^n \to T(\Sigma)$ were allowed. In this case probabilistic bisimilarity (Definition 5) may not be a congruence (Theorem 1). Example: PTSS (Σ, A, R), constants r, r', s in Σ, $A = \{a\}$, $R = \left\{ \frac{}{s \xrightarrow{a} \delta_s}, \frac{x \not\xrightarrow{a}}{g(x) \xrightarrow{a} f(\delta_x)} \right\}$ with $f(r) = r, f(r') = s$. Now $r \sim r'$ but $g(r) \not\sim g(r')$.

for $\theta = \sum_{i\in\{1,2\}} p_i\theta_i$. Case iii lifts the structural inductive construction of state terms to distribution terms. $\mathsf{DT}(\Sigma,\mathscr{D},\mathscr{V})$ is abbreviated as $\mathbb{DT}(\Sigma)$.

$\mathrm{MVar}\colon \mathbb{DT}(\Sigma) \to (\mathscr{V}\cup\mathscr{D}\to\mathbb{N})$ denotes for $\mathrm{MVar}(\theta)(\zeta)$ how often the variable ζ occurs in θ. For convex combinations $\sum_{i\in I} p_i\theta_i$ the maximal occurrence in some θ_i is considered because the probabilistic choice selects (probabilistically) exactly one of the summands. Formally, we have $\mathrm{MVar}(\mu)(\mu) = 1$, $\mathrm{MVar}(\mu)(\zeta) = 0$ if $\mu \neq \zeta$, $\mathrm{MVar}(\delta_t)(x) = \mathrm{MVar}(t)(x)$, $\mathrm{MVar}(\delta_t)(\mu) = 0$, $\mathrm{MVar}(\sum_{i\in I} p_i\theta_i)(\zeta) = \max_{i\in I}\mathrm{MVar}(\theta_i)(\zeta)$, and $\mathrm{MVar}(f(\theta_1,\ldots,\theta_{\mathsf{r}(f)}))(\zeta) = \sum_{i=1}^{\mathsf{r}(f)}\mathrm{MVar}(\theta_i)(\zeta)$. A substitution on state and distribution variables is a mapping $\sigma : (\mathscr{V}\cup\mathscr{D}) \to (\mathbb{T}(\Sigma)\cup\mathbb{DT}(\Sigma))$ such that $\sigma(x)\in\mathbb{T}(\Sigma)$ if $x\in\mathscr{V}$, and $\sigma(\mu)\in\mathbb{DT}(\Sigma)$ if $\mu\in\mathscr{D}$. A substitution extends to distribution terms by $\sigma(\delta_t) = \delta_{\sigma(t)}$, $\sigma(\sum_{i\in I} p_i\theta_i) = \sum_{i\in I} p_i\sigma(\theta_i)$ and $\sigma(f(\theta_1,\ldots,\theta_{\mathsf{r}(f)})) = f(\sigma(\theta_1),\ldots,\sigma(\theta_{\mathsf{r}(f)}))$. Notice that closed instances of distribution terms are probability distributions.

3 Probabilistic Transition System Specifications

Probabilistic transition systems (PTSs) generalize labelled transition systems (LTSs) by allowing for probabilistic choices in the transitions. We consider nondeterministic PTSs (Segala-type systems) [18] with countable state spaces.

Definition 1 (PTS)**.** *A* probabilistic labeled transition system *(PTS) is a triple $(T(\Sigma),A,\to)$, where Σ is a signature, A is a countable set of actions, and $\to \subseteq T(\Sigma)\times A\times\Delta(T(\Sigma))$ is a transition relation.*

We write $s \xrightarrow{a} \pi$ for $(s,a,\pi)\in\to$. PTSs are specified by means of transition system specifications [10, 11, 15, 17].

Definition 2 (Simple $nt\mu f\theta/nt\mu x\theta$-rule)**.** *A* simple $nt\mu f\theta$-rule *has the form:*

$$\frac{\{t_k \xrightarrow{a_k} \mu_k \mid k\in K\} \qquad \{t_l \xrightarrow{b_l}\!\!\!\!\!\!/\ \ \mid l\in L\}}{f(x_1,\ldots,x_{\mathsf{r}(f)}) \xrightarrow{a} \theta}$$

with $t_k,t_l\in\mathbb{T}(\Sigma)$, $a_k,b_l,a\in A$, $\mu_k\in\mathscr{D}$, $f\in\Sigma$, $x_1,\ldots,x_{\mathsf{r}(f)}\in\mathscr{V}$, $\theta\in\mathbb{DT}(\Sigma)$, and constraints:

1. *all μ_k for $k\in K$ are pairwise different;*
2. *all $x_1,\ldots,x_{r(f)}$ are pairwise different.*

A simple $nt\mu x\theta$-rule *is as above with source of its conclusion $x\in\mathscr{V}$ instead of $f(x_1,\ldots,x_{\mathsf{r}(f)})$. A* simple $nt\mu f\theta/nt\mu x\theta$-rule *is either a simple $nt\mu f\theta$-rule or a simple $nt\mu x\theta$-rule.*

The expressions $t_k \xrightarrow{a_k} \mu_k$ (resp. $t_l \xrightarrow{b_l}\!\!\!\!\!\!/\ \ $) above the line are called *positive* (resp. *negative*) *premises*. We call μ_k in $t_k \xrightarrow{a_k} \mu_k$ a *derivative* for each $x\in\mathrm{Var}(t_k)$. For rule ρ we denote the set of positive (resp. negative) premises by $\mathrm{pprem}(\rho)$ (resp. $\mathrm{nprem}(\rho)$), and the set of all premises by $\mathrm{prem}(\rho)$. A rule without premises is called an *axiom*. We allow the sets of positive and negative premises to be infinite. The expression $f(x_1,\ldots,x_{\mathsf{r}(f)}) \xrightarrow{a} \theta$ below the line is called *conclusion*, notation $\mathrm{conc}(\rho)$. The term $f(x_1,\ldots,x_{\mathsf{r}(f)})$ is called the *source* of ρ, notation $\mathrm{src}(\rho)$, and x_i are the *source variables*, notation $x_i\in\mathrm{src}(\rho)$. θ is the *target* of ρ, notation $\mathrm{trgt}(\rho)$. An expression $t \xrightarrow{a} \theta$ (resp. $t \xrightarrow{a}\!\!\!\!\!\!/\ \ $) is called a *positive* (resp. *negative*) *literal*. Hence, premises and conclusions are literals. We denote the set of variables in ρ by $\mathrm{Var}(\rho)$, *bound variables* by $\mathrm{bound}(\rho) = \{x_1,\ldots,x_{\mathsf{r}(f)}\}\cup\{\mu_k \mid k\in K\}$, and *free variables* by $\mathrm{free}(\rho) = \mathrm{Var}(\rho)\setminus\mathrm{bound}(\rho)$.

A *probabilistic transition system specification* (PTSS) in simple $nt\mu f\theta/nt\mu x\theta$-format, called simple $nt\mu f\theta/nt\mu x\theta$-PTSS for short, is a triple $P = (\Sigma,A,R)$ with Σ a signature, A a set of action labels, and R a set of simple $nt\mu f\theta/nt\mu x\theta$-rules.

As PTSS have negative premises, there are multiple approaches to assign a meaning (see [9] for an overview). We will use the stratification approach presented in [4] to assign to each PTSS $P = (\Sigma, A, R)$ (if possible) a PTS $(T(\Sigma), A, \rightarrow_P)$. A closed literal $t \xrightarrow{a} \pi$ (resp. $t \not\xrightarrow{a}$) *holds in* $\rightarrow_P$, notation $\rightarrow_P \models t \xrightarrow{a} \pi$ (resp. $\rightarrow_P \models t \not\xrightarrow{a}$), if $(t, a, \pi) \in \rightarrow_P$ (resp. there is no $\pi \in \Delta(T(\Sigma))$ s.t. $(t, a, \pi) \in \rightarrow_P$). A substitution σ extends to literals by $\sigma(t \xrightarrow{a} \mu) = \sigma(t) \xrightarrow{a} \sigma(\mu)$, and $\sigma(t \not\xrightarrow{a}) = \sigma(t) \not\xrightarrow{a}$, and to rules as expected.

Definition 3 (Stratification [4])**.** *Let $P = (\Sigma, A, R)$ be a PTSS. A function $S : T(\Sigma) \times A \times \Delta(T(\Sigma)) \rightarrow \alpha$, where α is an ordinal, is called a* stratification *of P if for every rule ρ*

$$\frac{\{t_k \xrightarrow{a_k} \mu_k \mid k \in K\} \quad \{t_l \not\xrightarrow{b_l} \mid l \in L\}}{f(x_1, \ldots, x_{r(f)}) \xrightarrow{a} \theta}$$

in R and substitution $\sigma : (\mathscr{V} \cup \mathscr{D}) \rightarrow (T(\Sigma) \cup \Delta(T(\Sigma)))$ we have: (i) $S(\sigma(t_k \xrightarrow{a_k} \mu_k)) \leq S(conc(\sigma(\rho)))$ for all $k \in K$, and (ii) $S(\sigma(t_l \xrightarrow{b_l} \mu)) < S(conc(\sigma(\rho)))$ for all $l \in L, \mu \in \mathscr{D}$. The set $S_\beta = \{\psi \mid S(\psi) = \beta\}$, with $\beta < \alpha$, is called a stratum*.*

We call *P stratifiable* if P has some stratification. A transition relation is constructed stratum by stratum in an increasing manner.

Definition 4 (Induced PTS [4])**.** *Let $P = (\Sigma, A, R)$ be a PTSS with stratification $S : T(\Sigma) \times A \times \Delta(T(\Sigma)) \rightarrow \alpha$. For all rules ρ, let $D(\rho)$ be the smallest regular cardinal greater than $|pprem(\rho)|$, and let $D(P)$ be the smallest regular cardinal such that $D(P) \geq D(\rho)$ for all $\rho \in R$. The* induced PTS *$(T(\Sigma), A, \rightarrow_{P,S})$ is defined by $\rightarrow_{P,S} = \bigcup_{\beta<\alpha} \rightarrow_{P_\beta}$, where $\rightarrow_{P_\beta} = \bigcup_{j \leq D(P)} \rightarrow_{P_{\beta,j}}$ and $\rightarrow_{P_{\beta,j}}$ is*

$$\rightarrow_{P_{\beta,j}} = \left\{ \psi \;\middle|\; \begin{array}{l} S(\psi) = \beta \text{ and } \exists \rho \in R \text{ and substitution } \sigma \text{ s.t. } \psi = conc(\sigma(\rho)), \text{ and} \\ (\bigcup_{\gamma<\beta} \rightarrow_{P_\gamma}) \cup (\bigcup_{j'<j} \rightarrow_{P_{\beta,j'}}) \models pprem(\sigma(\rho)), \text{ and} \\ (\bigcup_{\gamma<\beta} \rightarrow_{P_\gamma}) \models nprem(\sigma(\rho)) \end{array} \right\}$$

The induced PTS is independent from the chosen stratification [4]. We can construct for each simple *ntμfθ/ntμxθ*-PTSS (Σ, A, R) a PTSS (Σ, A, R') with only simple *ntμfθ*-rules that induces the same PTS [15]. The construction defines R' as R where each rule with a source of the form x is replaced by a set of rules where x is substituted by $f(x_1, \ldots, x_{r(f)})$ for each $f \in \Sigma$. Hence, all our results below for simple *ntμfθ*-PTSS generalize to simple *ntμfθ/ntμxθ*-PTSS.

Given a relation $\mathsf{R} \subseteq T(\Sigma) \times T(\Sigma)$, a set $X \subseteq T(\Sigma)$ is R-closed, denoted by R-closed(X), if $\mathsf{R}(X) \subseteq X$ where $\mathsf{R}(X) = \{y \in T(\Sigma) \mid \exists x \in X . x \,\mathsf{R}\, y\}$.

Definition 5 (Probabilistic Bisimulation [5, 14])**.** *Let $(T(\Sigma), A, \rightarrow)$ be a PTS. A symmetric relation $R \subseteq T(\Sigma) \times T(\Sigma)$ is a* probabilistic bisimulation *if whenever $t \,R\, t'$ and $t \xrightarrow{a} \pi$ then there exists a transition $t' \xrightarrow{a} \pi'$ such that $\pi \,R\, \pi'$, where*

$$\pi \,R\, \pi' \quad \textit{iff} \quad \textit{for all } X \subseteq T(\Sigma) \textit{ with } R\textit{-closed}(X) \textit{ we have } \pi(X) = \pi'(X).$$

Notice that this standard definition can be slightly reformulated to relate it to the later introduced ε-bisimulation (Definition 6) by requiring that $\pi \,\mathsf{R}\, \pi'$ iff $\pi(X) \leq \pi'(\mathsf{R}(X))$ for all $X \subseteq T(\Sigma)$ [7]. The union of all probabilistic bisimulations is the largest probabilistic bisimulation, called probabilistic bisimilarity, and denoted by $\sim$. We shall refer to probabilistic bisimulation as strict bisimulation to distinguish it from the later introduced relaxed notion of ε-bisimulation.

A crucial property of process description languages to ensure compositional modelling and verification is the compatibility of process operators with the behavioral relation chosen for the application

context. In algebraic terms the compatibility of a behavioral equivalence R with operator $f \in \Sigma$ is expressed by the congruence property which is defined as $f(t_1,\ldots,t_{\mathsf{r}(f)})\ \mathsf{R}\ f(u_1,\ldots,u_{\mathsf{r}(f)})$ whenever $t_i\ \mathsf{R}\ u_i$ for $i = 1,\ldots,\mathsf{r}(f)$. The rule format of Definition 2 is an instance of the $nt\mu f\theta/nt\mu x\theta$ rule format [15], which ensures that bisimilarity is a congruence.

Theorem 1 (Probabilistic Bisimilarity as a congruence [15])**.** *Let $P = (\Sigma, A, R)$ be a stratifiable simple $nt\mu f\theta/nt\mu x\theta$-PTSS. Then probabilistic bisimilarity is a congruence for all operators defined in P.*

In order to allow for robust reasoning on PTSs, the behavioral relations should allow for (limited) perturbation of probabilities [8]. ε-bisimulation is a behavioral relation based on strict probabilistic bisimulation, where the transfer condition is relaxed by some upper bound on the pertubation of probabilities.

Definition 6 (ε-Bisimulation [7])**.** *Let $(T(\Sigma), A, \to)$ be a PTS and $\varepsilon \in [0,1]$. A symmetric relation $R \subseteq T(\Sigma) \times T(\Sigma)$ is an* ε-bisimulation *if whenever $t\ R\ t'$ and $t \xrightarrow{a} \pi$ then there exists a transition $t' \xrightarrow{a} \pi'$ such that $\pi\ R\ \pi'$, where*

$$\pi\ R\ \pi' \quad \textit{iff} \quad \textit{for all } X \subseteq T(\Sigma) \textit{ we have } \pi(X) \leq \pi'(R(X)) + \varepsilon.$$

We call t and t' (resp. π and π') ε-bisimilar if $t\ \mathsf{R}\ t'$ (resp. $\pi\ \mathsf{R}\ \pi'$) for some ε-bisimulation R. Notice that ε-bisimulations are reflexive and symmetric but not necessarily transitive. ε-bisimulations are closed under union. We denote the largest ε-bisimulation, called ε-bisimilarity, by $\sim_\varepsilon$. According to [7], ε-bisimulations induce a pseudo-metric over the set of closed terms $d : T(\Sigma) \times T(\Sigma) \to [0,1]$ with $d(t,t') = \inf\{\varepsilon \in [0,1] \mid t \sim_\varepsilon t'\}$, where $\inf \emptyset = 1$. We say that t and t' are within the approximation bisimulation distance ε if $d(t,t') = \varepsilon$.

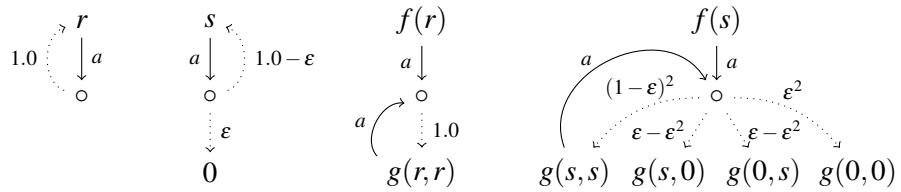
4 Expansivity of Process Combinators

The expansivity of an operator $f \in \Sigma$ is defined as the maximal approximate bisimulation distance of terms with an outermost function symbol f in relation to the approximate bisimulation distances of its arguments. In this section we quantify the expansivity of operators defined by a PTSS. We start by showing that the expansivity of an operator f defined by a rule ρ depends on (i) the *multiplicity* (i.e. number of occurrences) of source variables and their derivatives in the target of ρ; (ii) the *expansivity power* of operators (i.e. how much does the operator multiply the distance of its arguments) that define a context around the source variables or their derivatives; and (iii) the (reactive behavior) *discriminating power* of the premises of ρ.

Example 1 (Factors of Expansivity)**.** Let (Σ, A, R) be a PTSS with a signature Σ that contains constants $r, s, 0$, unary function symbols f, f_2, binary function symbols g, g_2, g_3 and a quaternary function symbol h, action set $A = \{a\}$, and axioms $R = \{r \xrightarrow{a} \delta_r, s \xrightarrow{a} [1-\varepsilon]\delta_s \oplus [\varepsilon]\delta_0\}$ for some fixed $\varepsilon \in (0,1)$. It is not hard to see that $d(r,s) = \varepsilon$ in the PTS induced by (Σ, A, R). Consider the rules:

$$\frac{x \xrightarrow{a} \mu}{f(x) \xrightarrow{a} g(\mu,\mu)} \qquad \frac{x_1 \xrightarrow{a} \mu_1 \quad x_2 \xrightarrow{a} \mu_2}{g(x_1,x_2) \xrightarrow{a} g(\mu_1,\mu_2)}$$

These rules together with R define R_2. In the first rule the derivative μ of source variable x appears twice in the rule target $g(\mu,\mu)$. The induced PTS of (Σ, A, R_2) contains the following transitions:



Observe that $d(f(r), f(s)) = 1-(1-\varepsilon)^2$. The power of 2 in the distance reflects directly the multiplicity of 2 of the derivative μ in the rule target. The same effect can be observed for multiple occurrences of source variables in the rule target, e.g. consider for the f-defining rule $g(\delta_x, \delta_x)$ instead of $g(\mu,\mu)$ as target.

Furthermore, the expansivity power of operators used in the rule target determine the expansivity of the operator defined by that rule. A simple example is the axiom $f_2(x) \xrightarrow{a} \delta_{f(x)}$. While the variable x occurs only once in the rule target, we still have $d(f_2(r), f_2(s)) = 1-(1-\varepsilon)^2$, because the operator f has an expansivity power of 2 wrt. its single argument. This indicates that the expansivity power of (arguments of) operators need to be defined recursively.

The multiplicity of source variables and their derivatives and the expansivity power of operators applied on those variables multiply. Consider the rules:

$$\frac{x_1 \xrightarrow{a} \mu_1 \quad x_2 \xrightarrow{a} \mu_2}{g(x_1,x_2) \xrightarrow{a} h(\mu_1,\mu_1,\mu_2,\mu_2)} \qquad \frac{\{x_i \xrightarrow{a} \mu_i \mid i = 1,\ldots,4\}}{h(x_1,x_2,x_3,x_4) \xrightarrow{a} h(\mu_1,\mu_2,\mu_3,\mu_4)}$$

These rules together with R_2 define R_3. Now $d(f(r), f(s)) = 1-(1-\varepsilon)^4$. As explained above for R_2, in the rule defining operator f the derivative μ appears twice in the rule target. Additionally the operator g that is applied to μ has for both of its arguments an expansivity power of two because in the g-defining rule the derivatives μ_1, μ_2 of both arguments x_1, x_2 appear twice in the rule target.

The expansivity power of an operator may be unbounded. Consider the recursive unary operator f defined by the rules:

$$\frac{x \xrightarrow{a} \mu}{f(x) \xrightarrow{a} g(f(\mu), f(\mu))} \qquad \frac{x_1 \xrightarrow{a} \mu_1 \quad x_2 \xrightarrow{a} \mu_2}{g(x_1,x_2) \xrightarrow{a} g(\mu_1,\mu_2)}$$

These rules together with R define R_4. In the rule that defines the operator f the derivative μ occurs twice in the target. Moreover, each occurrence of μ is put in the context of that operator f, which is defined by this rule (recursive call). Additionally both occurrences of $f(\mu)$ are put in the binary context g, which enforces that the distances of the two copies of μ multiply. Recursive multiplication of the distances leads to an approximate bisimulation distance of $d(f(r), f(s)) = 1$. The expansivity power of f will in this case be denoted by ∞.

On the other hand, an operator may also absorb the approximate bisimulation distance. Consider the rules:

$$\frac{x \xrightarrow{a} \mu}{f(x) \xrightarrow{a} g_2(\mu,\mu)} \qquad \frac{x \xrightarrow{a} \mu}{f_2(x) \xrightarrow{a} g_3(\mu,\mu)} \qquad \frac{}{g_3(x_1,x_2) \xrightarrow{a} \delta_0}$$

These rules together with R define R_5. The first rule applies the undefined operator g_2 to the two copies of the derivative μ. As g_2 has no rules, we get $d(f(r), f(s)) = 0$. Similarly, the rule defining f_2 applies operator g_3 in the target. The operator g_3 allows one to derive an unconditional move to the idle process 0. Hence, $d(f_2(r), f_2(s)) = 0$.

However, if the reactive behavior of the process associated to a source variable is tested by some premise, then the operator defined by this rule may discriminate states with different reactive behavior. Consider the rules:

$$\frac{x \xrightarrow{a} \mu}{f(x) \xrightarrow{a} g(\mu,\mu)} \qquad \frac{x_1 \xrightarrow{a} \mu_1 \quad x_2 \xrightarrow{a} \mu_2}{g(x_1,x_2) \xrightarrow{a} \delta_0}$$

These rules together with R define R_6. We get $d(f(r),f(s)) = 1-(1-\varepsilon)^2$ because the a-transition of term r leads to a distribution where all states can perform the action a, but the a-transition of term s leads to a distribution where only states with a total probability mass of $1-\varepsilon$ can perform the action a. ■

We denote by R_f those rules of R that define the operator f. We define by $\nabla(f,i) \in \{0,1\}$ the (reactive behavior) discriminating power of argument i of f. Formally, $\nabla(f,i) = 1$ if the source variable x_i appears in a premise of some $\rho \in R_f$, i.e., if for some $\rho \in R_f$ there is a $t_k \xrightarrow{a_k} \mu_k \in \mathrm{pprem}(\rho)$ with $x_i \in \mathrm{Var}(t_k)$ or a $t_l \xrightarrow{b_l} \not\rightarrow \in \mathrm{nprem}(\rho)$ with $x_i \in \mathrm{Var}(t_l)$. Otherwise, $\nabla(f,i) = 0$. With $\mathbb{N}^\infty$ we denote $\mathbb{N} \cup \{\infty\}$, with the natural ordering extended by $n < \infty$ for each $n \in \mathbb{N}$, and the usual arithmetic extended for summation by $\infty + n = n + \infty = \infty + \infty = \infty$ for $n \geq 0$ and multiplication by $0 \cdot \infty = \infty \cdot 0 = 0$ and $n \cdot \infty = \infty \cdot n = \infty \cdot \infty = \infty$ for $n \geq 1$.

We quantify the expansivity power of operators $f \in \Sigma$ as least fixed point of a monotone function. Let (Σ, A, R) with $\Sigma = (F, \mathsf{r})$ be a PTSS. We define a poset $\mathscr{S} = (S, \sqsubseteq)$ with $S = S_F \times S_T$, $S_F = F \times \mathbb{N} \to \mathbb{N}^\infty$, $S_T = (\mathbb{T}(\Sigma) \cup \mathbb{DT}(\Sigma)) \to ((\mathscr{V} \cup \mathscr{D}) \to \mathbb{N}^\infty)$, equipped with the point-wise partial order $(m_F, m_T) \sqsubseteq (m'_F, m'_T)$ iff $m_F(f,i) \leq m'_F(f,i)$, for all $f \in F, i \in \mathbb{N}$, and $m_T(t)(\zeta) \leq m'_T(t)(\zeta)$ for all $t \in \mathbb{T}(\Sigma) \cup \mathbb{DT}(\Sigma), \zeta \in \mathscr{V} \cup \mathscr{D}$. Elements of S are pairs of maps (m_F, m_T). $m_F(f,i)$ denotes the expansivity power of argument i in operator f, i.e., how much the operator f multiplies the approximate bisimulation distance of argument i. $m_T(t)(\zeta)$ defines the frequency of variable $\zeta \in \mathscr{V} \cup \mathscr{D}$ in the state or distribution term $t \in \mathbb{T}(\Sigma) \cup \mathbb{DT}(\Sigma)$ weighted by the expansivity power of the operators applied on top of ζ. $\mathscr{S}$ forms a complete lattice with bottom element $\bot$ and top element $\top$, defined by constant maps $\bot((f,i),(t,\zeta)) = (0,0)$ and $\top((f,i),(t,\zeta)) = (\infty,\infty)$ for each $f \in F$, $i \in \mathbb{N}$, $t \in \mathbb{T}(\Sigma) \cup \mathbb{DT}(\Sigma)$, $\zeta \in \mathscr{V} \cup \mathscr{D}$.

Proposition 1 *$\mathscr{S}$ is a complete lattice.*

The function $M : S \to S$ defined in Fig. 1 computes in parallel the expansivity power of arguments of operators, and the multiplicities of variables in terms weighted by the expansivity power of the operators applied on top of them. The expansivity power $m'_F(f,i)$ of argument i of operator f is defined as the maximum expansivity power over each f-defining rule $\rho \in R_f$. For $\rho \in R_f$ the expansivity power is defined as the sum of the multiplicity of x_i in the rule target $\mathrm{trgt}(\rho)$ and of the multiplicity of x_i in some premise $t_k \xrightarrow{a} \mu_k \in \mathrm{pprem}(\rho)$ weighted by the multiplicity of the derivative μ_k in the rule target $\mathrm{trgt}(\rho)$. Note that source variables and derivatives in the rule target contribute equally to the expansivity power of an argument. The multiplicity $m'_T(t)(\zeta)$ of ζ in a state term t counts the occurrences of variable ζ in t and weights them by the expansivity power of the operators applied on top of ζ. The multiplicity $m'_T(\theta)(\zeta)$ of ζ in a distribution term θ counts the occurrences of variable ζ in θ and weights them by the expansivity power of the operators applied on top of ζ, but at least by the discriminating power of those operators. Note that the discriminating power of operators is considered only for distribution terms. To understand this, consider the reactive behavior of ε-bisimilar state and distribution terms. For a state term $f(t_1,\ldots,t_{\mathsf{r}(f)})$ we have that $\sigma(t_i) \sim_{\varepsilon_i} \sigma'(t_i)$ implies $\sigma(t_i) \xrightarrow{a}$ iff $\sigma(t'_i) \xrightarrow{a}$ for each $a \in A$, i.e., $\sigma(t_i)$ and $\sigma(t'_i)$ agree on their immediate reactive behavior. However, for a distribution term $f(\theta_1,\ldots,\theta_{\mathsf{r}(f)})$ we have that if $\sigma(\theta_i) \sim_{\varepsilon_i} \sigma'(\theta_i)$ then $\sigma(\theta_i)$ and $\sigma'(\theta_i)$ may have states with different reactive behavior (cf. R_6 in Example 1).

Function $M : S \to S$ is defined by $M(m_F, m_T) = (m'_F, m'_T)$ with

$$m'_F(f,i) = \sup_{\rho \in R_f} \left(m_T(\mathrm{trgt}(\rho))(x_i) + \sum_{\substack{a_k \\ t_k \xrightarrow{} \mu_k \in \\ \mathrm{pprem}(\rho)}} m_T(t_k)(x_i) \cdot m_T(\mathrm{trgt}(\rho))(\mu_k) \right)$$

$$m'_T(t)(\zeta) = \begin{cases} 1 & \text{if } \zeta \in \mathcal{V} \text{ and } t = \zeta \\ \sum_{i=1}^{\mathsf{r}(f)} (m_F(f,i) \cdot m_T(t_i)(\zeta)) & \text{if } t = f(t_1, \ldots, t_{\mathsf{r}(f)}) \\ 0 & \text{otherwise} \end{cases}$$

$$m'_T(\theta)(\zeta) = \begin{cases} 1 & \text{if } \theta = \zeta \\ m_T(t)(\zeta) & \text{if } \theta = \delta_t \\ \max_{i \in I} \left(m_T(\theta_i)(\zeta) \right) & \text{if } \theta = \sum_{i \in I} p_i \theta_i \\ \sum_{i=1}^{\mathsf{r}(f)} (\max(m_F(f,i), \overline{\gamma}(f,i)) \cdot m_T(\theta_i)(\zeta)) & \text{if } \theta = f(\theta_1, \ldots, \theta_{\mathsf{r}(f)}) \\ 0 & \text{otherwise} \end{cases}$$

Figure 1: Function to quantify the approximate bisimulation multiplicity

M is order-preserving. This ensures the existence and uniqueness of the least fixed point of M by the Knaster-Tarski fixed point theorem.

Proposition 2 *M is order-preserving.*

We denote the least fixed point of M by (ω_F, ω_T). We call $\omega_F(f,i)$ the expansivity power of argument i of operator f, and $\omega_T(t)(\zeta)$ the weighted multiplicity of variable ζ in term t. The expansivity power of f allows us to derive an upper bound on the approximate bisimulation distance between terms $f(t_1, \ldots, t_{\mathsf{r}(f)})$ and $f(t'_1, \ldots, t'_{\mathsf{r}(f)})$ expressed in relation to the approximate bisimulation distances ε_i between the arguments t_i and t'_i.

Definition 7 (Expansivity bound)**.** *The* expansivity bound $\exp^f$ *of operator $f \in \Sigma$ wrt. the approximate bisimulation distances ε_i of its arguments $i = 1, \ldots, \mathsf{r}(f)$ is defined by*

$$\exp^f(\varepsilon_1, \ldots, \varepsilon_{\mathsf{r}(f)}) = 1 - \prod_{i=1}^{\mathsf{r}(f)} (1 - \varepsilon_i)^{\omega_F(f,i)}$$

Notice that $\exp^f(\varepsilon_1, \ldots, \varepsilon_{\mathsf{r}(f)}) = 0$ if $\varepsilon_i = 0$ for all arguments i with $\omega_F(f,i) > 0$. In particular, we have $\exp^f(\varepsilon_1, \ldots, \varepsilon_{\mathsf{r}(f)}) = 0$ if all $\varepsilon_i = 0$. We call an argument i of operator $f \in \Sigma$ (behavioral distance) *absorbing* if $\omega_F(f,i) = 0$.

We demonstrate first the application of the expansivity bound and prove later its correctness.

Example 1 (continued)**.** For the PTSS (Σ, A, R_2) we have $\omega_F(f,1) = 2$ because $\omega_F(g,1) = \omega_F(g,2) = 1$. Terms r and s with approximate bisimulation distance $d(r,s) = \varepsilon$ agree by $1 - \varepsilon$ on their behavior.

Thus, the pair of processes (r,r) and (s,s) agree by $(1-\varepsilon)^2$ on their behavior. Hence, they disagree by $1-(1-\varepsilon)^2$ on their behavior. This gives a behavioral distance of $d(f(r),f(s)) = 1-(1-\varepsilon)^2$.

We continue with PTSS (Σ,A,R_3). For operator h we have $\omega_F(h,1) = \omega_F(h,2) = \omega_F(h,3) = \omega_F(h,4) = 1$, for g we have $\omega_F(g,1) = \omega_F(g,2) = 2$ and thus for f we get $\omega_F(f,1) = 4$. For PTSS (Σ,A,R_4) the recursive definition of f applied to the two occurrences of the derivative μ in the rule target gives $\omega_F(f,1) = \infty$. The (behavioral distance) absorbing effect of f and f_2 in (Σ,A,R_5) results in $\omega_F(f,1) = \omega_F(f_2,1) = \omega_F(g_2,1) = \omega_F(g_2,2) = \omega_F(g_3,1) = \omega_F(g_3,2) = 0$. In (Σ,A,R_6) the (reactive behavior) discriminating power $\overline{\gamma}(g,1) = \overline{\gamma}(g,2) = 1$ of operator g leads to $\omega_F(f,1) = 2$. ■

Now we can show that the approximate bisimulation distance between terms $f(t_1,\ldots,t_{\mathsf{r}(f)})$ and $f(t'_1,\ldots,t'_{\mathsf{r}(f)})$ is bounded by the expansivity bound.

Theorem 2 (Expansivity bound of simple $nt\mu f\theta$-PTSS)**.** *Let (Σ,A,R) be a stratifiable simple $nt\mu f\theta$-PTSS. Then for each operator $f \in \Sigma$ we have*

$$f(t_1,\ldots,t_{\mathsf{r}(f)}) \sim_\varepsilon f(t'_1,\ldots,t'_{\mathsf{r}(f)}) \quad \textit{whenever} \quad t_i \sim_{\varepsilon_i} t'_i \ \textit{for } i = 1,\ldots,\mathsf{r}(f)$$

with $\varepsilon = \exp^f(\varepsilon_1,\ldots,\varepsilon_{\mathsf{r}(f)})$.

Theorem 2 implies Theorem 1 by considering $\varepsilon_i = 0$ for all $i = 1,\ldots,\mathsf{r}(f)$ and exploiting that $\sim_0$ is in fact the strict probabilistic bisimilarity.

Our target was to define the expansivity power $\omega_F(f)(i)$ of the argument i of operator $f \in \Sigma$ in order to characterize the behavioral distance of terms with outermost function symbol f. We conclude this section by outlining how the expansivity bound could be further refined. Sequential composition $\_;\_$ is defined by the following rules [4]

$$\frac{x \xrightarrow{a} \mu}{x;y \xrightarrow{a} \mu;\delta_y}\ a \neq \surd \qquad\qquad \frac{x \xrightarrow{\surd} \mu \quad y \xrightarrow{a} \mu'}{x;y \xrightarrow{a} \mu'}$$

Action $\surd$ denotes successful termination. The expansivity power $\omega_F(;)(1) = \omega_F(;)(2) = 1$ gives an expansivity bound $\exp^{;}(\varepsilon_1,\varepsilon_2) = 1-(1-\varepsilon_1)(1-\varepsilon_2)$. However, the sequential composition describes separate moves of either process x or process y. Hence, the expansivity of $\_;\_$ is actually bounded by $1-\min(1-\varepsilon_1,1-\varepsilon_2)$. In general, if multiple rules define an operator f, then the expansivity power and weighted multiplicity should be quantified per rule instead of per operator. In detail, the expansivity power $\omega_F(f)(i)$ should take a rule ρ instead of f as argument, and the weighted multiplicity $\omega_T(t,x)$ should take a tree of rules instead of term t as argument. We leave this as future work.

5 Specification of Non-expansive Process Combinators

Non-expansivity is the quantitative analogue of the congruence property of (strict) probabilistic bisimulation. Intuitively, non-expansivity means that different processes are not more different when they are put in the same context.

Definition 8 (Non-expansivity)**.** *Let $(T(\Sigma),A,\to_P)$ be the PTS induced by the PTSS $P = (\Sigma,A,R)$. An operator $f \in \Sigma$ is* non-expansive *if*

$$f(t_1,\ldots,t_{\mathsf{r}(f)}) \sim_\varepsilon f(t'_1,\ldots,t'_{\mathsf{r}(f)}) \quad \textit{whenever} \quad t_i \sim_{\varepsilon_i} t'_i \ \textit{for all } i = 1,\ldots,\mathsf{r}(f)$$

with $\varepsilon = \min(\sum_{i=1}^{\mathsf{r}(f)} \varepsilon_i, 1)$.

We call f *expansive* if f is not non-expansive. Argumentation for this linear upper bound and a discussion on alternative upper bounds like maximum norm or Euclidean norm can be found in [20].

From the expansivity bound $\exp^f$ (Definition 7) of operator $f \in \Sigma$ it follows that f is non-expansive if $\omega_F(f,i) \leq 1$ for all $i = 1,\ldots,\mathsf{r}(f)$. This yields the following rule format.

Definition 9 (ε-*ntμfθ* rule format)**.** *A simple ntμfθ-rule ρ is an ε-ntμfθ*-rule *if for each $x_i \in \mathrm{src}(\rho)$ we have*

$$\mathrm{MVar}(\mathrm{trgt}(\rho))(x_i) + \sum_{\substack{a_k \\ t_k \xrightarrow{} \mu_k \in \\ pprem(\rho)}} \mathrm{MVar}(t_k)(x_i) \cdot \mathrm{MVar}(\mathrm{trgt}(\rho))(\mu_k) \leq 1.$$

A PTSS $P = (\Sigma, A, R)$ is in ε-ntμfθ format, ε-ntμfθ-PTSS for short, if all rules in R are in the ε-ntμfθ rule format.

Theorem 3 (Non-expansivity of ε-*ntμfθ*-PTSS)**.** *Let (Σ, A, R) be a stratifiable ε-ntμfθ-PTSS. Then all operators $f \in \Sigma$ are non-expansive.*

The constraints of the ε-*ntμfθ* rule format are easy to verify. It suffices to count the occurrences of source variables and derivatives in the rule target. There is no need for recursive reasoning over other rules. We deliberately decided against the (slightly more general) rule format which could be given as simple *ntμfθ*-rules ρ that define some operator $f \in \Sigma$ and for which the only requirement would be $\omega_F(f,i) \leq 1$ for all $i = 1,\ldots,\mathsf{r}(f)$. We justify this by considering the extension of a PTSS $P = (\Sigma, A, R)$ to $P' = (\Sigma, A, R')$ with $R \subseteq R'$. If P is in ε-*ntμfθ* format, then in order to decide if P' is in ε-*ntμfθ* format only the rules in $R' \setminus R$ need to be verified wrt. the ε-*ntμfθ* format constraints. On the contrary, the generalized rule format would require that whenever a rule is added, all other rules are again validated with respect to the format constraints. For instance, consider the set of rules R_6 in Example 1. The rule defining operator f alone would be non-expansive. However, by adding the rule defining operator g (even though g is non-expansive), operator f becomes expansive.

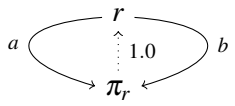
6 Applications

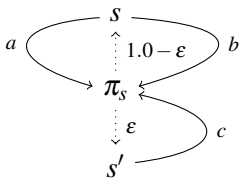
The standard process combinators sequential composition, (probabilistic and non-probabilistic) choice, and (probabilistic and non-probabilistic) CCS and CSP like parallel composition [2, 4] are all in the ε-*ntμfθ*-format. On the other hand, recursion and iteration operators may be expansive if they replicate (some of) their arguments. We consider the replication operator of π-calculus. The nondeterministic variant $!\_$ and the probabilistic variant $!^p\_$ with $p \in (0,1) \cap \mathbb{Q}$ are defined by the rules [16]:

$$\frac{x \xrightarrow{a} \mu}{!x \xrightarrow{a} \mu \parallel \delta_{!x}} \qquad \frac{x \xrightarrow{a} \mu}{!^p x \xrightarrow{a} \mu \oplus_p (\mu \parallel \delta_{!^p x})} \qquad \frac{x_1 \xrightarrow{a} \mu_1 \quad x_2 \xrightarrow{a} \mu_2}{x_1 \parallel x_2 \xrightarrow{a} \mu_1 \parallel \mu_2}$$

The first two rules defining both variants of the replication operator are not in ε-*ntμfθ*-format. The expansivity power of both operators is unbounded with $\omega_F(!)(1) = \omega_F(!^p)(1) = \infty$. Hence, both operators are expansive. However, if the synchronous parallel composition defined in the third rule above is replaced by the non-communicating asynchronous parallel composition, then both variants of the replication operator would become non-expansive.

We summarize the structural patterns of rules that may lead to expansive behavior in Table 1. None of these rules is in the ε-*ntμfθ* format. For cases 1 to 7 the expansivity power of f is $\omega_F(f)(1) = 2$ and, therefore, the expansivity bound is $\exp^f(\varepsilon) = d(f(r), f(s)) = 1 - (1-\varepsilon)^2$. Cases 8 and 9 indicate that





Common rule

$$\frac{x_1 \xrightarrow{a} \mu_1 \quad x_2 \xrightarrow{a} \mu_2}{x_1 \parallel x_2 \xrightarrow{a} \mu_1 \parallel \mu_2}$$

	Description	Rule
1	Non-linearity of the rule target wrt. a source variable	$\frac{x \xrightarrow{a} \mu}{f(x) \xrightarrow{a} \delta_x \parallel \delta_x}$
2	Non-linearity of the rule target wrt. a derivative	$\frac{x \xrightarrow{a} \mu}{f(x) \xrightarrow{a} \mu \parallel \mu}$
3	Non-linearity of a state term in the rule target	$\frac{x \xrightarrow{a} \mu}{f(x) \xrightarrow{a} \delta_{x \parallel x}}$
4	Non-linearity of a term in a premise	$\frac{x \parallel x \xrightarrow{a} \mu}{f(x) \xrightarrow{a} \mu}$
5	Multiple derivatives of a source variable in the rule target	$\frac{x \xrightarrow{a} \mu_1 \quad x \xrightarrow{b} \mu_2}{f(x) \xrightarrow{a} \mu_1 \parallel \mu_2}$
6	Source and derivative in the rule target	$\frac{x \xrightarrow{a} \mu}{f(x) \xrightarrow{a} \delta_x \parallel \mu}$
7	Multiple derivatives weighted by convex combination in rule target	$\frac{x \xrightarrow{a} \mu}{f(x) \xrightarrow{a} [0.5]\mu \parallel \mu \oplus [0.5]\delta_{s'}}$
8	Lookahead by existential test in quantitative premise	$\frac{x \xrightarrow{a} \mu_1 \quad \mu_1(y) > 0 \quad y \xrightarrow{c} \mu_2}{f(x) \xrightarrow{a} \mu_2}$
9	Lookahead by universal test in quantitative premise	$\frac{x \xrightarrow{a} \mu \quad \mu(Y) \geq 1 \quad \{y \xrightarrow{a} \mu_y \mid y \in Y\}}{f(x) \xrightarrow{a} \delta_r}$

Table 1: SOS rules that specify expansive operators

lookahead cannot be admitted and we need to employ simple $nt\mu f\theta$-rules (Definition 2) instead of $nt\mu f\theta$-rules [15]. The expressions $\mu_1(y) > 0$ and $\mu_1(Y) \geq 1$ (with $y \in \mathscr{V}$, $Y \subseteq \mathscr{V}$) are quantitative premises as introduced by the $nt\mu f\theta/nt\mu x\theta$ format [4]. As argued above, ε-bisimilar instances may have states with different reactive behavior. For instance, in case 8, while distributions π_r and π_s are ε-bisimilar, only π_s has in its support a state that can perform a c-move. Similarly, in case 9, only for π_r we have that all states in the support can perform an a-move.

We conjecture that a notion of *non-expansivity up to* ε for some $\varepsilon \in [0,1]$ (bounded non-expansivity) would allow for limited lookahead. An operator is non-expansive up to ε if it is non-expansive whenever

its arguments have an approximate bisimulation distance of at most ε. In this case, quantitative premises $\mu(Y) \geq p$ that measure the probability of Y and test against the boundary p could be allowed, if p is in the interval $p \in [\varepsilon, 1-\varepsilon]$. On the other hand, quantitative premises with $p < \varepsilon$ or $p > 1-\varepsilon$ cannot be permitted because they allow for lookahead with respect to probabilistic choices that do not mimic each other's reactive behavior.

The expansivity bound (Definition 7) allows rule formats to be derived for alternative compositionality requirements. For instance, consider an n-ary process combinator $\otimes$ with the compositionality requirements that the approximate bisimulation distance of the combined processes should not depend on the approximate bisimulation distance of processes at some argument $i \in \{1,\ldots,n\}$. From the expansivity bound we derive that either argument i of operator $\otimes$ is behavioral distance absorbing ($\omega_F(\otimes)(i) = 0$), or the application context guarantees that processes for argument i are strictly bisimilar ($\varepsilon_i = 0$).

The non-expansivity requirement (Definition 8) is in fact the Manhattan norm, or more general the p-norm $(\sum_{i=1}^{\mathsf{r}(\otimes)} \varepsilon_i^p)^{1/p}$ with $p = 1$. Consider the alternative compositionality requirement that the expansivity of a process combinator $\otimes$ should be bounded by the p-norm with $p > 1$ (which includes the Euclidean norm by $p = 2$ and the maximum norm by $p \to \infty$). From the expansivity bound we derive that $\omega_F(\otimes)(i) = 1$ for at most one argument i, and all other arguments $j \neq i$ are behavioral distance absorbing with $\omega_F(\otimes)(j) = 0$.

7 Conclusion and Future Work

We studied structural specifications of probabilistic processes that are robust with respect to bounded implementation and measurement errors of probabilistic behavior. We provided for each process combinator an upper bound on the distance between the combined processes using the structural specification of the process combinator (Theorem 2). We derived an appropriate rule format that guarantees non-expansivity (standard compositionality requirement) of process combinators (Theorem 3). All standard process algebraic operators are compatible for approximate reasoning and satisfy the rule format, except operators which replicate processes and combine them by synchronous parallel composition. We exemplified how rule formats for non-standard compositionality requirements can be derived.

Our work can be extended in several directions. In Section 4 and Section 6 we sketched already how the expansivity bound can be further refined and how a restricted form of lookahead in the rules specifying the process combinators could be admitted. The techniques and results developed in this paper for approximate bisimulation can be carried over to bisimulation metrics. Initial work in this direction suggests that the ε-$nt\mu f\theta$ format presented in this paper ensures also non-expansivity for the bisimulation metric based on the Kantorovich and Hausdorff metric. Moreover, for the bisimulation metric the rule format can be further generalized because in this case convex combinations weigh the distance and multiplicity of processes (unlike approximate bisimilarity, see case 7 of Table 1). Furthermore, we will investigate the expansivity of process combinators and rule formats for variants of bisimulation metrics and ε-bisimulation that discount the influence of future transitions [6, 22].

Acknowledgements We are grateful to Josée Desharnais for discussions on ε-bisimulation, Matteo Mio for discussions on approximate semantics of structurally defined probabilistic systems, and Wan Fokkink and David Williams for feedback on earlier versions of this paper. Furthermore, we thank the anonymous referees for thorough reviews and very helpful comments.

References

[1] Giorgio Bacci & Marino Miculan (2012): *Structural Operational Semantics for Continuous State Probabilistic Processes*. In: *Proc. CMCS'12*, *LNCS* 7399, Springer, pp. 71–89, doi:10.1007/978-3-642-32784-1_5.

[2] Falk Bartels (2004): *On Generalised Coinduction and Probabilistic Specification Formats*. Ph.D. thesis, VU University Amsterdam.

[3] Franck van Breugel & James Worrell (2005): *A Behavioural Pseudometric for Probabilistic Transition Systems*. *Theor. Comput. Sci.* 331(1), pp. 115–142, doi:10.1016/j.tcs.2004.09.035.

[4] Pedro R. D'Argenio & Matias David Lee (2012): *Probabilistic Transition System Specification: Congruence and Full Abstraction of Bisimulation*. In: *Proc. FoSSaCS'12*, *LNCS* 7213, Springer, pp. 452–466, doi:10.1007/978-3-642-28729-9_30.

[5] Josée Desharnais, Vineet Gupta, Radha Jagadeesan & Prakash Panangaden (2003): *Approximating Labelled Markov Processes*. *Information and Computation* 184(1), pp. 160 – 200, doi:10.1016/S0890-5401(03)00051-8.

[6] Josée Desharnais, Vineet Gupta, Radha Jagadeesan & Prakash Panangaden (2004): *Metrics for Labelled Markov Processes*. *Theor. Comput. Sci.* 318(3), pp. 323–354, doi:10.1016/j.tcs.2003.09.013.

[7] Josée Desharnais, Francois Laviolette & Mathieu Tracol (2008): *Approximate Analysis of Probabilistic Processes: Logic, Simulation and Games*. In: *Proc. QEST'08*, IEEE, pp. 264–273, doi:10.1109/QEST.2008.42.

[8] Alessandro Giacalone, Chi-Chang Jou & Scott A. Smolka (1990): *Algebraic Reasoning for Probabilistic Concurrent Systems*. In: *Proc. IFIP TC2 Working Conf. on Prog. Concepts and Methods*, pp. 443–458.

[9] Rob J. van Glabbeek (2004): *The Meaning of Negative Premises in Transition System Specifications II*. *J. Log. Algebr. Program.* 60-61, pp. 229–258, doi:10.1016/j.jlap.2004.03.007.

[10] Jan Friso Groote (1993): *Transition System Specifications with Negative Premises*. *Theor. Comput. Sci.* 118(2), pp. 263–299, doi:10.1016/0304-3975(93)90111-6.

[11] Jan Friso Groote & Frits Vaandrager (1992): *Structured Operational Semantics and Bisimulation as a Congruence*. *Inf. Comput.* 100, pp. 202–260, doi:10.1016/0890-5401(92)90013-6.

[12] Ruggero Lanotte & Simone Tini (2005): *Probabilistic Congruence for Semistochastic Generative Processes*. In: *Proc. FoSSaCS'05*, *LNCS* 3441, Springer, pp. 63–78, doi:10.1007/978-3-540-31982-5_4.

[13] Ruggero Lanotte & Simone Tini (2009): *Probabilistic Bisimulation as a Congruence*. *ACM TOCL* 10, pp. 1–48, doi:10.1145/1462179.1462181.

[14] Kim G. Larsen & Arne Skou (1991): *Bisimulation Through Probabilistic Testing*. *Inf. Comput.* 94, pp. 1–28, doi:10.1016/0890-5401(91)90030-6.

[15] Matias David Lee, Daniel Gebler & Pedro R. D'Argenio (2012): *Tree Rules in Probabilistic Transition System Specifications with Negative and Quantitative Premises*. In: *Proc. EXPRESS/SOS'12*, *EPTCS* 89, pp. 115–130, doi:10.4204/EPTCS.89.9.

[16] Matteo Mio & Alex Simpson (2013): *A Proof System for Compositional Verification of Probabilistic Concurrent Processes*. In: *Proc. FoSSaCS'13*, *LNCS* 7794, Springer, pp. 161–176, doi:10.1007/978-3-642-37075-5_11.

[17] Gordon Plotkin (1981): *A Structural Approach to Operational Semantics*. Report DAIMI FN-19, Aarhus University.

[18] Roberto Segala (1995): *Modeling and Verification of Randomized Distributed Real-Time Systems*. Ph.D. thesis, MIT.

[19] Simone Tini (2008): *Non Expansive ε-bisimulations*. In: *Proc. AMAST'08*, *LNCS* 5140, Springer, pp. 362–376, doi:10.1007/978-3-540-79980-1_27.

[20] Simone Tini (2010): *Non-expansive ε-bisimulations for Probabilistic Processes*. *Theoret. Comput. Sci.* 411, pp. 2202–2222, doi:10.1016/j.tcs.2010.01.027.

[21] Mathieu Tracol (2010): *Approximate Verification of Probabilistic Systems*. Dissertation, LRI, Université Paris-Sud.

[22] Mathieu Tracol, Josée Desharnais & Abir Zhioua (2011): *Computing Distances between Probabilistic Automata*. In: *Proc. QAPL'11*, *EPTCS* 57, pp. 148–162, doi:10.4204/EPTCS.57.11.