Tree rules in probabilistic transition system specifications with negative and quantitative premises*

Matias David Lee† Daniel Gebler‡ Pedro R. D'Argenio†

†FaMAF – CONICET
Universidad Nacional de Córdoba
Ciudad Universitaria, X5000HUA Córdoba
Argentina
{lee,dargenio}@famaf.unc.edu.ar

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

Probabilistic transition system specifications (PTSSs) in the *ntμfv/ntμxv* format provide structural operational semantics for Segala-type systems that exhibit both probabilistic and nondeterministic behavior and guarantee that bisimilarity is a congruence. Similar to the nondeterministic case of the rule format *tyft/tyxt*, we show that the well-foundedness requirement is unnecessary in the probabilistic setting. To achieve this, we first define a generalized version of the *ntμfv/ntμxv* format in which quantitative premises and conclusions include nested convex combinations of distributions. Also this format guarantees that bisimilarity is a congruence. Then, for a given (possibly non-well-founded) PTSS in the new format, we construct an equivalent well-founded PTSS consisting of only rules of the simpler (well-founded) probabilistic ntree format. Furthermore, we develop a proof-theoretic notion for these PTSSs that coincides with the existing stratification-based meaning in case the PTSS is stratifiable. This continues the line of research lifting structural operational semantic results from the nondeterministic setting to systems with both probabilistic and nondeterministic behavior.

1 Introduction

Plotkin's structural operational semantics [19] is a popular method to provide a rigorous interpretation to specification and programming languages. The interpretation is given in terms of transition systems. The method has been formalized with an algebraic flavor as *transition systems specifications (TSS)* [5,6,13,14, etc.]. Basically, a TSS contains a signature, a set of labels, and a set of rules. The signature defines the terms in the language. Labels represent actions performed by a process (i.e., a term over the signature) in one step of the execution (i.e., one transition). Rules define how a process should behave (i.e., produce a transition) in terms of the behavior of its subprocesses. That is, rules define compositionally the transition system associated to each term of the language. This technique has been widely studied mainly on the realm of languages and process algebras describing only non-deterministic behavior (see [18] for an overview).

The introduction of probabilistic process algebras [1, 12, etc.] motivated the need for a theory of structural operational semantics to define *probabilistic* transition systems. A few results have appeared in this direction, notably [2, 3, 7, 15, 16]. All these works introduced rule formats that ensures that bisimulation equivalence is a congruence for operators whose semantics is defined within such format. The most general of those formats is the *ntμfv/ntμxv* format [7] that provides semantics in terms of Segala's probabilistic automata [20].

*Supported by Project ANPCYT PAE-PICT 02272, SeCyT-UNC, Eramus Mundus Action 2 Lot 13A EU Mobility Programme 2010-2401/001-001-EMA2 and EU 7FP grant agreement 295261 (MEALS).

B. Luttik and M.A. Reniers (Eds.): Combined Workshop on Expressiveness in Concurrency and Structural Operational Semantics (EXPRESS/SOS 2012)
EPTCS 89, 2012, pp. 115–130, doi:10.4204/EPTCS.89.9

© M. D. Lee, D. Gebler & P. R. D'Argenio
This work is licensed under the
Creative Commons Attribution License.

The *ntμfv/ntμxν* format is the probabilistic relative to the *ntyft/ntyxt* format [13] extending it in two ways. First, it is designed to deal with probabilistic transitions of the form $t \xrightarrow{a} \pi$, where t is a term in the appropriate signature, and π is a distribution on terms. Second, it includes quantitative premises that allow for probabilistic testing of the form $\pi(\{t_1,\ldots,t_n\}) > q$, that is, it allows to verify if the probability that the system moves to one state (i.e. term) in $\{t_1,\ldots,t_n\}$ according to π is greater than $q \in [0,1]$. The congruence theorem for the *ntμfv/ntμxν* format [7, Thm. 12] states that if a probabilistic transition system specification (PTSS) P has all its rules in *ntμfv/ntμxν* format, then bisimulation equivalence is a congruence for all operators in P. Unfortunately, [7] missed an important condition: rules have to be well-founded (basically, there should not be a cyclic dependency on the terms appearing in the premises of the rule). This paper will correct this mistake.

The well-foundedness condition has also appeared from the very beginning in the non-deterministic setting. Most of the formats have it implicit as they did not allowed lookahead. Congruence theorems for formats with lookahead such as *tyft/tyxt* [14] or *ntyft/ntyxt* [13] explicitly demanded TSS to be well-founded. It remained unknown for a while whether such condition was actually required until Fokkink and van Glabbeek proved it unnecessary [9]. The proof proceeds by reducing a TSS in *tyft/tyxt* format (not necessarily well-founded) to an equivalent TSS containing only so called *tree rules* (i.e., well-founded rules in *tyft* format with premises containing only variables instead of arbitrary open terms). Similarly, they showed that a TSS in *ntyft/ntyxt* format can be translated into an equivalent TSS containing only ntree rules (tree rules with negative premises which are not necessarily restricted to single variables).

In this paper, we also show that the restriction to well-founded PTSSs is not necessary to guarantee congruence. We also proceed by reducing a PTSS in *ntμfv/ntμxν* format to an equivalent PTSS containing only *pntree rules*. However, a pntree rule cannot simply be defined as an *ntμfv* rule where positive premises are restricted to the form $x \xrightarrow{a} \mu$, with x and μ being term and distribution variables, respectively. It turns out that quantitative premises in *ntμfv/ntμxν* rules are too limited. The *ntμfv/ntμxν* format only allows for quantitative premises of the form $\mu(Y) \trianglerighteq q$ with μ being a distribution variable, Y an infinite set of term variables, $\trianglerighteq \in \{>,\geq\}$, and $q \in [0,1]$. Instead, the pntree format requires premises of the form $\theta(Y) \trianglerighteq q$ where θ is a nested convex combinations of products of distribution variables. We call these objects *distribution terms*. So, we extend the *ntμfv/ntμxν* format to deal with distribution terms, and prove, more generally, that a PTSS in the new format — called *ntμfθ/ntμxθ* — can be translated into an equivalent PTSS with only pntree rules (hence, well-founded). Just like for the case of the *ntyft/ntyxt* format, full negative premises are required, i.e., negative premises in pntree rules cannot be limited to the form $x \overset{a}{\not\rightarrow}$, with x being a term variable.

Summarizing, the following results are introduced in this paper:

- We define the *ntμfθ/ntμxθ* format, which extends the *ntμfv/ntμxν* format to deal with distribution terms in quantitative premises.
- We prove that if a PTSS is in *ntμfθ/ntμxθ* format and it is well-founded, then bisimulation equivalence is a congruence for all its operators. This also corrects the mistake in the proof of Theorem 12 in [7] which omitted to consider the well-foundedness hypothesis.
- We show that for all PTSS in *ntμfθ/ntμxθ* format (not necessarily well-founded) there is a PTSS with only pntree rules that defines exactly the same probabilistic transition relation (by "defines" we mean "has as a *supported model*")
- We dropped the well-foundedness hypothesis from the congruence theorem: since every pntree rule is also a well-founded *ntμfθ* rule, the previous results imply that bisimulation equivalence is a congruence for all operators of a (not necessarily well-founded) PTSS in *ntμfθ/ntμxθ* format.

- Besides, in the process, we also redefined important concepts for PTSS originally defined for TSS, in particular, the concept of "well supported proof".

2 Preliminaries

We assume the presence of an infinite set of (term) variables $\mathcal{V}$ and we let $x, y, z, x', x_0, x_1, \ldots$ range over $\mathcal{V}$. A *signature* is a structure $\Sigma = (F, \mathsf{r})$, where (i) F is a set of *function names* disjoint with $\mathcal{V}$, and (ii) $\mathsf{r} : F \to \mathbb{N}_0$ is a *rank function* which gives the arity of a function name; if $f \in F$ and $\mathsf{r}(f) = 0$ then f is called a *constant name*. Let $W \subseteq \mathcal{V}$ be a set of variables. The set of Σ-terms over W, notation $T(\Sigma, W)$ is the least set satisfying: (i) $W \subseteq T(\Sigma, W)$, and (ii) if $f \in F$ 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 abbreviated as $T(\Sigma)$; the elements of $T(\Sigma)$ are called *closed terms*. $T(\Sigma, \mathcal{V})$ is abbreviated as $\mathbb{T}(\Sigma)$; the elements of $\mathbb{T}(\Sigma)$ are called *open terms*. $Var(t) \subseteq \mathcal{V}$ is the set of variables in the open term t.

In order to deal with languages that describe probabilistic behavior we need expressions denoting probability distributions. Let $\Delta(T(\Sigma))$ denote the set of all (discrete) probability distributions on $T(\Sigma)$. We let $\pi, \pi', \pi_0, \pi_1, \ldots$ range over $\Delta(T(\Sigma))$. As usual, for $\pi \in \Delta(T(\Sigma))$ and $T \subseteq T(\Sigma)$, we define $\pi(T) = \sum_{t \in T} \pi(t)$. For $t \in T(\Sigma)$, let δ_t denote the Dirac distribution, i.e., $\delta_t(t) = 1$ and $\delta_t(t') = 0$ if $t \neq t'$. Moreover, the product measure $\prod_{i=1}^n \pi_i$ is defined by $(\prod_{i=1}^n \pi_i)(t_1, \ldots, t_n) = \prod_{i=1}^n \pi_i(t_i)$. In particular, if $n = 0$, $(\prod_{j \in \emptyset} \pi_j) = \delta_{()}$ is the distribution that assigns probability 1 to the empty tuple. Let $g : T(\Sigma)^n \to T(\Sigma)$ and recall that $g^{-1}(t') = \{\vec{t} \in T(\Sigma)^n \mid g(\vec{t}) = t'\}$. Then $(\prod_{i=1}^n \pi_i) \circ g^{-1}$ is a well defined probability distribution on closed terms. In particular, if $g : T(\Sigma)^0 \to T(\Sigma)$ and $g(()) = t$, then $(\prod_{j \in \emptyset} \pi_j) \circ g^{-1} = \delta_{()} \circ g^{-1} = \delta_t$.

For a term $t \in \mathbb{T}(\Sigma)$ we let δ_t be an *instantiable Dirac distribution*. That is, δ_t is a symbol that takes value $\delta_{t'}$ when variables in t are substituted so that t becomes a closed term $t' \in T(\Sigma)$. Let $\mathcal{D} = \{\delta_t : t \in \mathbb{T}(\Sigma)\}$ be the set of instantiable Dirac distributions. A *distribution variable* is a variable that takes values on $\Delta(T(\Sigma))$. Let $\mathcal{M}$ be an infinite set of distribution variables. Let $\mu, \mu', \mu_0, \mu_1, \ldots$ range over $\mathcal{M}$ and $\zeta, \zeta', \zeta_0, \zeta_1, \ldots$ range over $\mathcal{M} \cup \mathcal{V}$. Let $D \subseteq \mathcal{M}$ be a set of distribution variables and $V \subseteq \mathcal{V}$ be a set of term variables. The set of *distribution terms* over D and V, notation $\mathsf{DT}(\Sigma, D, V)$ is the least set satisfying: (i) $D \cup \{\delta_t : t \in T(\Sigma, V)\} \subseteq \mathsf{DT}(\Sigma, D, V)$, and (ii) $\sum_{i \in I} p_i (\prod_{n_i \in N_i} \theta_{n_i}) \circ g_i^{-1} \in \mathsf{DT}(\Sigma, D, V)$ where $p_i \in (0, 1]$ with $\sum_{i \in I} p_i = 1$, each g_i is a function s.t. $g_i : T(\Sigma)^{N_i} \to T(\Sigma)$, and $\theta_{n_i} \in \mathsf{DT}(\Sigma, D, V)$. Intuitively, $g_i^{-1}(t)$ decomposes term t into its sub-terms $t_1, \ldots, t_{N_i}$ and probability $\theta(t)$ of term t is calculated as the convex combination of the product probability of its sub-terms $\theta_1(t_i), \ldots, \theta_{N_i}(t_{N_i})$. $\mathsf{DT}(\Sigma, \emptyset, \emptyset)$ is abbreviated as $\mathsf{DT}(\Sigma)$; the elements of $\mathsf{DT}(\Sigma)$ are actual distributions on terms. $\mathsf{DT}(\Sigma, \mathcal{M}, \mathcal{V})$ is abbreviated as $\mathbb{DT}(\Sigma)$. $Var(\theta) \subseteq \mathcal{M} \cup \mathcal{V}$ is the set of (distribution and term) variables appearing in θ.

A substitution is a mapping that assigns terms to variables. In our case we need to extend this notion to distribution terms and instantiable Dirac distributions. A *substitution* ρ is a mapping in $(\mathcal{V} \cup \mathcal{M}) \to (\mathbb{T}(\Sigma) \cup \mathbb{DT}(\Sigma))$ such that $\rho(x) \in \mathbb{T}(\Sigma)$ whenever $x \in \mathcal{V}$, and $\rho(\mu) \in \mathbb{DT}(\Sigma)$ whenever $\mu \in \mathcal{M}$. A substitution ρ extends to open terms and sets of terms as usual, to instantiable Dirac distributions by $\rho(\delta_t) = \delta_{\rho(t)}$ and to distribution terms by $\rho(\sum_{i \in I} p_i (\prod_{n_i \in N_i} \theta_{n_i}) \circ g_i^{-1}) = \sum_{i \in I} p_i (\prod_{n_i \in N_i} \rho(\theta_{n_i})) \circ g_i^{-1}$. Notice that the construction of distribution terms ensures that closed substitution instances of distribution terms denote indeed probability distribution.

3 Probabilistic Transition System Specifications

A (probabilistic) transition relation describes the behavior of a process by prescribing the possible actions it can perform at each state. Each action is described with a label on the relation and the evolution to

the next state is given by a probability distribution on terms. We will follow the probabilistic automata style of [20] which generalize the so called reactive model [17]. Let Σ be a signature and A be a set of labels. A *transition relation* is a set $\rightarrow \subseteq PTr(\Sigma, A)$, where $PTr(\Sigma, A) = T(\Sigma) \times A \times \Delta(T(\Sigma))$. We denote $(t, a, \pi) \in \rightarrow$ by $t \xrightarrow{a} \pi$.

Transition relations are usually defined by means of structured operational semantics in Plotkin's style [19]. We follow the approach of [6, 13, 14] which provides an algebraic characterization for transition system specifications.

Definition 1. *A* probabilistic transition system specification *(PTSS) is a triple* $P = (\Sigma, A, R)$ *where* $\Sigma = (F, r)$ *is a signature, A is a set of labels, and R is a set of rules of the form:*

$$\frac{\{t_k \xrightarrow{a_k} \mu_k : k \in K\} \cup \{t_l \xrightarrow{b_l} \not\rightarrow : l \in L\} \cup \{\theta_j(W_j) \gtrless_j q_j : j \in J\}}{t \xrightarrow{a} \theta}$$

where K, L, J *are index sets,* $t, t_k, t_l \in \mathbb{T}(\Sigma)$*,* $a, a_k, b_l \in A$*,* $\mu_k \in \mathcal{M}$*,* $W_j \subseteq \mathcal{V}$*,* $\gtrless_j \in \{>, \geq, <, \leq\}$*,* $q_j \in [0, 1]$ *and* $\theta_j, \theta \in \mathbb{DT}(\Sigma)$

An expression of the form $t \xrightarrow{a} \theta$, (resp. $t \xrightarrow{a} \not\rightarrow$, $\theta(W) \gtrless p$) is a *positive literal* (resp. *negative literal, quantitative literal*) where $t \in \mathbb{T}(\Sigma)$, $a \in A$, $\theta \in \mathbb{DT}(\Sigma)$, $W \subseteq Var \cup T(\Sigma)$ and $p \in [0, 1]$. For any rule $r \in R$, literals above the line are called *premises*, notation prem(r); the literal below the line is called *conclusion*, notation conc(r). We denote with pprem(r) (nprem(r), qprem(r)) the set of positive (negative, quantitative) literals of the rule r. A rule r is called *positive* if nprem(r) $= \emptyset$. A PTSS is called positive if it has only positive rules. A rule r without premises is called an *axiom*. In general, we allow the sets of positive, negative, and quantitative premises to be infinite.

Substitutions provide instances to the rules of a PTSS that, together with some appropriate machinery, allows us to define probabilistic transition relations. Given a substitution ρ, it extends to literals as follows: $\rho(t \xrightarrow{a} \not\rightarrow) = \rho(t) \xrightarrow{a} \not\rightarrow$, $\rho(\theta(W) \gtrless p) = \rho(\theta)(\rho(W)) \gtrless p$, and $\rho(t \xrightarrow{a} \theta) = \rho(t) \xrightarrow{a} \rho(\theta)$. Then, the notion of substitution extends to rules as expected. We say that r' is a (closed) instance of a rule r if there is a (closed) substitution ρ so that $r' = \rho(r)$.

We say that ρ is a *proper substitution of* r if for all quantitative premises $\rho(\theta(W) \gtrless p)$ of r it holds that $\rho(\theta(w)) > 0$ for all $w \in W$. Thus, if ρ is proper, all terms in $\rho(W)$ are in the support of $\rho(\theta)$. Proper substitutions avoid the introduction of spurious terms. This is of particular importance for the conservative extension theorem of [7, Theorem 14]. We use only this kind of substitution in the paper.

As has already been argued many times (e.g. [6, 11, 13]), transition system specifications with negative premises do not uniquely define a transition relation and different reasonable techniques may lead to incomparable models. In any case, we expect that a transition relation associated to a PTSS P (i) respects the rules of P, that is, whenever the premises of a closed instance of a rule of P belong to the transition relation, so does its conclusion; and (ii) it does not include more transitions than those explicitly justified, i.e., a transition is defined only if it is the conclusion of a closed rule whose premises are in the transition relation. The first notion corresponds to that of model, and the second one to that of supported transition.

Before formally defining these notions we introduce some notation. Given a transition relation $\rightarrow \subseteq PTr(\Sigma, A)$, a positive literal $t \xrightarrow{a} \pi$ *holds in* $\rightarrow$, notation $\rightarrow \models t \xrightarrow{a} \pi$, if $(t, a, \pi) \in \rightarrow$. A negative literal $t \xrightarrow{a} \not\rightarrow$ *holds in* $\rightarrow$, notation $\rightarrow \models t \xrightarrow{a} \not\rightarrow$, if there is no $\pi \in \Delta(T(\Sigma))$ s.t. $(t, a, \pi) \in \rightarrow$. A quantitative literal $\pi(T) \gtrless p$ *holds in* $\rightarrow$, notation $\rightarrow \models \pi(T) \gtrless p$ precisely when $\pi(T) \gtrless p$. Notice that the satisfaction of a quantitative literal does not depend on the transition relation. We nonetheless use this last notation as it turns out to be convenient. Given a set of literals H, we write $\rightarrow \models H$ if $\forall \phi \in H : \rightarrow \models \phi$.

Definition 2. *Let $P = (\Sigma, A, R)$ be a PTSS. Let $\to \subseteq PTr(\Sigma, A)$ be a probabilistic transition system (PTS). Then $\to$ is* a supported model of *P if it satisfies that: $\psi \in \to$ iff there is a rule $\frac{H}{\chi} \in R$ and a proper substitution ρ s.t. $\rho(\chi) = \psi$ and $\to \models \rho(H)$. For $\to$ to be a* model *of P we only require that the "if" holds, and for $\to$ to be* supported by *P we only require that the "only if" holds.*

We have already pointed out that PTSSs with negative premises do not uniquely define a transition relation. In fact, a PTSS may have more than one supported model. For instance, the PTSS with the single constant f, set of labels $\{a, b\}$ and the two rules $\frac{f \xrightarrow{a} \mu}{f \xrightarrow{a} \delta_f}$ and $\frac{f \not\xrightarrow{a}}{f \xrightarrow{b} \delta_f}$, has two supported models: $\{f \xrightarrow{a} \delta_f\}$ and $\{f \xrightarrow{b} \delta_f\}$. We will not dwell on this problem which has been studied at length in [6] and [11] in a non-probabilistic setting. Instead we present two different approaches to resolve this problem: stratification and well supported proofs.

3.1 Stratification

A stratification defines an order on closed positive literals that ensures that the validity of a transition does not depend on the negation of the same transition.

Definition 3. *Let $P = (\Sigma, A, R)$ be a PTSS. A function $S : PTr(\Sigma, A) \to \alpha$, where α is an ordinal, is called a* stratification *of P (and P is said to be* stratified*) if for every rule*

$$r = \frac{\{t_k \xrightarrow{a_k} \mu_k : k \in K\} \cup \{t_l \not\xrightarrow{b_l} : l \in L\} \cup \{\theta_j(W_j) \gtrless q_j : j \in J\}}{t \xrightarrow{a} \theta}$$

and proper substitution $\rho : (\mathcal{V} \cup \mathcal{M}) \to (T(\Sigma) \cup \Delta(T(\Sigma)))$ it holds that: (i) for all $k \in K$, $S(\rho(t_k \xrightarrow{a_k} \mu_k)) \leq S(conc(r))$, and (ii) for all $l \in L$ and $\mu \in \mathcal{M}$, $S(\rho(t_l \xrightarrow{b_l} \mu)) < S(conc(r))$. Each set $S_\beta = \{\phi \mid S(\phi) = \beta\}$, with $\beta < \alpha$, is called a stratum*. If for all $k \in K$, $S(\rho(t_k \xrightarrow{a_k} \mu_k)) < S(conc(r))$, then the stratification is said to be* strict*.*

A transition relation is constructed stratum by stratum in an increasing manner by transfinite recursion. If it has been decided whether a transition in a stratum $S_{\beta'}$, with $\beta' < \beta$, is valid or not, we already know the validity of the negative premise occurring in the premises of a transition φ in stratum S_β (since all positive instances of the negative premises are in strictly lesser strata) and hence we can determine the validity of φ. Notice that a stratification does not take quantitative premises into account because their satisfaction does not depend on the transition relation.

Definition 4. *Let $P = (\Sigma, A, R)$ be a PTSS with a stratification $S : PTr(\Sigma, A) \to \alpha$ for some ordinal α. For all rules r, let $\mathsf{D}(r)$ be the smallest regular cardinal such that $\mathsf{D}(r) \geq |pprem(r)|$, and let $\mathsf{D}(P)$ be the smallest regular cardinal such that $\mathsf{D}(P) \geq \mathsf{D}(r)$ for all $r \in R$. The transition relation $\to_{P,S}$* associated with *P (and based on S) is defined by $\to_{P,S} = \bigcup_{\beta<\alpha} \to_{P_\beta}$, where each $\to_{P_\beta} = \bigcup_{j \leq \mathsf{D}(P)} \to_{P_{\beta,j}}$ and each $\to_{P_{\beta,j}}$ is defined by*

$$\to_{P_{\beta,j}} = \Big\{ \psi \;\Big|\; S(\psi) = \beta \text{ and } \exists r \in R \text{ and proper substitution } \rho \text{ s.t. } \psi = conc(\rho(r)),\ (\textstyle\bigcup_{\gamma<\beta} \to_{P_\gamma}) \cup (\bigcup_{j'<j} \to_{P_{\beta,j'}}) \models qprem(\rho(r)) \cup pprem(\rho(r)) \text{ and } (\textstyle\bigcup_{\gamma<\beta} \to_{P_\gamma}) \models nprem(\rho(r)) \Big\}$$

A PTSS P with rules $R = \left\{ \frac{f \xrightarrow{a} \mu}{f \xrightarrow{a} \delta_f}, \frac{f \not\xrightarrow{a}}{f \xrightarrow{b} \delta_f} \right\}$ can be stratified by $S(f \xrightarrow{a} \delta_f) = 0$ and $S(f \xrightarrow{b} \delta_f) = 1$. This stratification induces the transition relation $\to_{P,S} = \{f \xrightarrow{b} \delta_f\}$. Because (non-strict) stratifications allow

that positive premises are in the same stratum as the conclusion, the validity of a premise may depend on a rule with a conclusion literal of the same stratum. In this case, the construction of $\rightarrow_{P_\beta}$ requires to iterate up to $D(P)$ times, denoted by $\bigcup_{j\leq \mathsf{D}(P)} \rightarrow_{P_{\beta,j}}$, to decide the the validity of all literals of this stratum.

The existence of a stratification guarantees the existence of a supported model. In fact, such model is the one in Def. 4 (Theorem 1). Furthermore, all stratification define the same supported model (Theorem 2) which allows to omit the stratification symbol in $\rightarrow_{P,S}$ and use $\rightarrow_P$ instead. Moreover, strict stratification ensures uniqueness of the supported model (Theorem 3). The proofs follow closely their non-probabilistic counterparts in [13] (Theorem 2.15, Lemma 2.16 and Theorem 2.18, resp.). The only actual difference lies on the quantitative premises, which do not pose any particular problem since their validity depends only on the substitution.

Theorem 1. *Let P be a PTSS with stratification S. Then $\rightarrow_{P,S}$ is a supported model of P.*

Theorem 2. *Let P be a PTSS. For all stratifications S, S' of P it holds $\rightarrow_{P,S} = \rightarrow_{P,S'}$.*

Theorem 3. *Let P be a PTSS with a strict stratification S. Then $\rightarrow_{P,S}$ is the only supported model of P.*

3.2 Proof structures

In this section we introduce the notion of *provable rules from a PTSS*. To define this notion we use *proof structures* [9]. A proof structure is like a derivation tree where the rules do not share variable names. The connection between the conclusion of a rule r and a premise ψ in other rule is represented by a mapping ϕ from rules to literals, i.e. $\phi(r) = \psi$. A substitution *matches* with a proof structure if both the conclusion and the premise related by ϕ are mapped to the same literal. Thus, matching substitutions translate a proof structure into an actual derivation tree. As a consequence, a matching substitution applied to a proof structure defines a *provable rule* in which the premises are the leaves of the derivation tree and the conclusion is the root. The absence of shared variables allows to define substitution on proof structures avoiding name clashes. Provable rules will be used in the following way through the paper: given a PTSS P we take the set of provable rules from P with a particular format, these rules will be used to define a a new PTSS P', then we show that P and P' derive the same PTS.

A PTSS is *small* if for each of its rules the cardinality of its collection of premises does not exceed the cardinality of the set of variables V. Small PTSS ensure that there are enough variables to construct the proof structures.

Definition 5. *A* proof structure *is a tuple $\langle B, r, \phi\rangle$ such that*

- *$r \in B$ and B is a set of transition rules which do not have any variables in common,*
- *ϕ is an injective mapping from $B \setminus \{r\}$ to the collection of positive premises in B, such that each chain $b_0, b_1, \ldots$ in B, with $\phi(b_{i+1})$ is a premise of b_i, is a finite chain.*

Let $top(B, r, \phi)$ be the set of all premises of rules in B that are outside the image of ϕ. Let $qtop(B, r, \phi)$ be the set of all quantitative premises in $top(B, r, \phi)$.

We introduce a partial well-order $<$ on proof structures to allow inductive reasoning. Define the partial order $<$ by $(B', r', \phi') < (B, r, \phi)$ iff $B' \subset B$, ϕ' is ϕ restricted to $B' \setminus \{r'\}$, $\mathrm{top}(B', r', \phi') \subseteq \mathrm{top}(B, r, \phi)$, and there is a chain $b_0, b_1, \ldots, b_n$ with $b_0 = r$, $b_n = r'$, $n > 0$ and $\phi(b_{i+1})$ is a premise of b_i.

A substitution σ *matches* with the proof structure (B, r, ϕ) if $\sigma(\mathrm{conc}(b)) = \sigma(\phi(b))$ for every $b \in B \setminus \{r\}$.

Definition 6. *Let $H = H_p \cup H_n \cup H_q$ a set of literals s.t. H_p, (resp. H_n and H_q) is a set of positive (resp. negative and open quantitative) literals. A rule $\frac{H}{c}$ is provable from a small PTSS $P = (\Sigma, A, R)$, notation $P \vdash \frac{H}{c}$, if $c \in H$ or there is a proof structure (B, r, ϕ) such that each rule in B is in R modulo α-conversion and there is a substitution σ that matches with (B, r, ϕ) such that:*

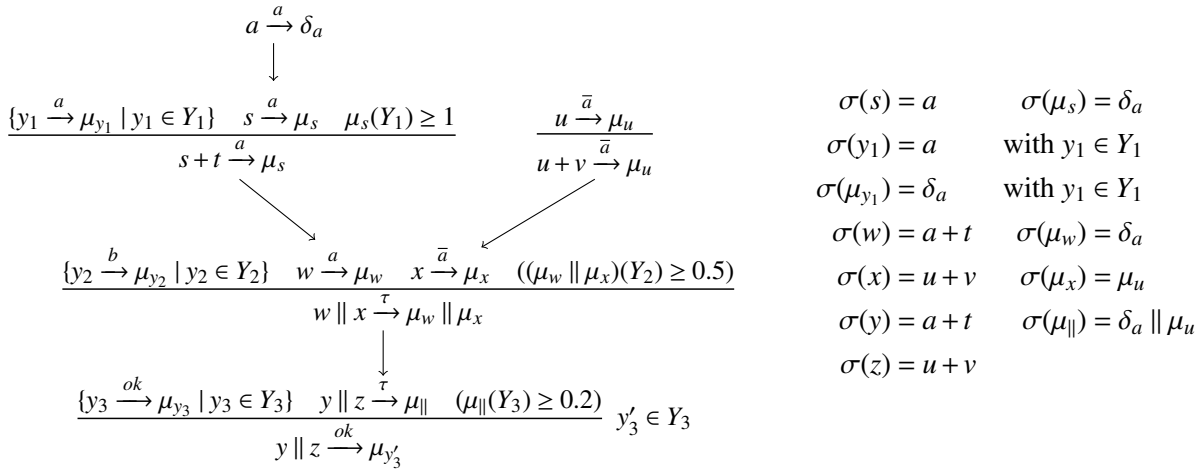

Figure 1: An example of proof structure. (See Example 1)

- $\sigma(top(B,r,\phi) - qtop(B,r,\phi)) \subseteq H$,
- *if* $\psi \in \sigma(qtop(B,r,\phi))$ *is a closed quantitative premise then* ψ *holds, otherwise* $\psi \in H_q$ *and*
- $\sigma(conc(r)) = c$.

Note that closed quantitative literals do not need to be included in the premise of a provable rule because their validity can be decided without further instantiation. Notice additionally that all negative literals of premises of rules in B are included in H and thus no negative literals can be derived.

Example 1. *Let* $P = \langle \Sigma, A, R \rangle$ *be a PTSS with* $\{a, +, \|\} \subseteq \Sigma$, $\{a, \overline{a}, b, \tau, ok\} \subseteq A$ *and all rules in Fig. 1 appear in R. Let* (B, r, ϕ) *the proof structure of Figure 1 where mapping* ϕ *is represented by the arrows. Let* σ *be the substitution defined in Fig. 1, with* $\sigma(\zeta) = \zeta$ *for any other (term or distribution) variable not specified in the figure. Then the following rule is provable from P:*

$$\frac{u \xrightarrow{\overline{a}} \mu_u \quad \{y_2 \xrightarrow{b} \mu_{y_2} \mid y_2 \in Y_2\} \quad ((\delta_a \parallel \mu_u)(Y_2) \geq 0.5) \quad \{y_3 \xrightarrow{ok} \mu_{y_3} \mid y_3 \in Y_3\} \quad ((\delta_a \parallel \mu_u)(Y_3) \geq 0.2)}{(a+t) \parallel (u+v) \xrightarrow{ok} \mu_{y_3'}} \quad (1)$$

Both in Fig. 1 and in the above rule we used shorthand notations for the different distribution terms. We write $(\mu_w \parallel \mu_x)$ *and* $(\delta_a \parallel \mu_u)$ *instead of* $(\mu_w \times \mu_x) \circ \|^{-1}$ *and* $((\delta_{()} \circ k_a^{-1}) \times \mu_x) \circ \|^{-1}$, *with* $k_a(()) = a$, *respectively (trivial summations are omitted).*

Since $\sigma(y_1) = a$ *for all* $y_1 \in y_1$ *and* $\sigma(\mu_s) = \delta_a$, *then* $\sigma(\mu_s(Y_1) \geq 1) = (\delta_a(\{a\}) \geq 1)$ *is closed, and moreover, it holds. As a consequence, it does not appear as a premise of rule (1). Also notice that* $\mu_\parallel$ *was substituted by* $(\delta_a \parallel \mu_u)$. *This is why we needed to upgrade the format of [7] to consider the more complex distribution terms on the quantitative premises instead of only distribution variables.*

The set of all provable rules from a PTSS can be alternatively defined in a recursive manner without using the notion of proof structure (Def. 7). We prove that both definitions are equivalent in Lemma 1.

Definition 7. *The* provable closure *of a PTSS* $P = \langle \Sigma, A, R \rangle$ *is the smallest set* $R^\vdash$ *of rules such that*

- *if* $c \in H$ *then* $\frac{H}{c} \in R^\vdash$,
- *if* $r \in R$ *and there is a substitution* σ *such that*

– *for all* $p \in pprem(r) \cup nprem(r)$ *it holds* $\frac{H}{\sigma(p)} \in R^{\vdash}$ *and*
– *for all* $p \in qprem(r)$ *if* $\sigma(p)$ *is not a closed literal then* $\frac{H}{\sigma(p)} \in R^{\vdash}$*, otherwise* $\sigma(p)$ *holds*

then $\frac{H}{\sigma(conc(r))} \in R^{\vdash}$.

Lemma 1. *A rule* $\frac{H}{c}$ *is provable from a small PTSS* $P = \langle \Sigma, A, R \rangle$ *iff* $\frac{H}{c} \in R^{\vdash}$.

The following lemma is an immediate consequence of Def. 7.

Lemma 2. *Let* P *and* P' *be two PTSS such that all rules in* P' *are provable from* P*. Then all rules provable from* P' *are also provable from* P*.*

3.3 Well-supported proofs

In the following we adapt the notion of *well-supported proof* [11] to PTSS. In the following, we say that literals $t \xrightarrow{a} \pi$ and $t \not\xrightarrow{a}$ *deny each other*.

Definition 8. *A* well-supported proof *of a closed literal* ψ *from a PTSS* $P = (\Sigma, A, R)$ *is a well-founded, upwardly branching tree of which the nodes are labeled by positive or negative literals, such that*

- *the root is labeled by* ψ*, and*
- *if* χ *is the label of the node* q *and* $\{\chi_k \mid k \in K\}$ *is the set of labels of the nodes directly above* q*, then:*
  - *if* χ *is a positive literal then there is a rule* $r \in R$ *and a closed proper substitution* ρ *such that* $\{\chi_k \mid k \in K\} = pprem(\rho(r)) \cup nprem(\rho(r))$*, the quantitative premises* $qprem(\rho(r))$ *are valid and* $conc(\rho(r)) = \chi$*,*
  - *if* χ *is a negative premise then for all* $P \vdash \frac{N}{\phi}$ *with* ϕ *a closed literal denying* χ*, a literal in* $\{\chi_k \mid k \in K\}$ *denies a literal in* N*.*

A literal ψ *is* ws-provable*, notation* $P \vdash_{ws} \psi$*, if there is a well-supported proof of* ψ *from* P*. A literal* ψ *is* ws-refutable *if there is a literal* ψ' *ws-provable from* P *and* ψ *denies* ψ'*.*

Notice that nodes in the proof tree of Def. 8 are not quantitative literals. This is due to the fact that the validity of closed quantitative literals is already known. In fact, the definition requires that all quantitative literal introduced by a rule r should become valid after substitution.

We say that a PTSS P is *complete* if for all closed literal $t \not\xrightarrow{a}$, $P \vdash_{ws} t \xrightarrow{a} \pi$ for some distribution π or $P \vdash_{ws} t \not\xrightarrow{a}$. In addition, P is *consistent* if there are no pair of literals derived from p that deny each other. We will focus only on complete PTSSs. The transition relation based on well-supported proofs associated to a (complete) PTSS P (denoted by $\rightarrow_{ws}$) is the set of ws-provable transitions of P.

Lemma 3. *Let* P *be a PTSS. If* P *is complete then it is also consistent.*

Lemma 3 allows us to show that, for any stratifiable PTSS, the model obtained using well-supported proofs coincides with the model obtained through stratification. Notice that this does not imply that the methods are equivalent: it could be the case that a PTSS is complete but not stratifiable (see [11, Prop. 27]).

Lemma 4. *Let* P *be a PTSS with stratification* S *and* ψ *a positive or negative literal, then* $\psi \in \rightarrow_{ws}$ *iff* $\rightarrow_{P,S} \models \psi$*.*

The proof of this lemma follows the same structure of its non-probabilistic counterpart (see [11, Prop. 25]).

The next lemma states that it suffices to show that the same rules having only negative premises are provable in two different PTSSs to state that these PTSSs define the same set of ws-provable transitions.

Lemma 5. *Let* P *and* P' *be two PTSSs over the same signature such that* $P \vdash \frac{H}{c}$ *iff* $P' \vdash \frac{H}{c}$ *for all closed rule* $\frac{H}{c}$ *with* H *containing only negative premises. Then* $P \vdash_{ws} \psi$ *iff* $P' \vdash_{ws} \psi$ *for all closed literal* ψ*.*

4 The $nt\mu f\theta/nt\mu x\theta$ format

In this section we revise the *ntμfv/ntμxv* format of [7] adapting it to the richer quantitative premises introduced before. Furthermore we correct some mistakes of [7].

Before, we recall the notion of bisimulation on PTSs [17]. Given a relation $\mathsf{R} \subseteq T(\Sigma)\times T(\Sigma)$, a set $Q \subseteq T(\Sigma)$ is R-closed if for all $t \in Q$ and $t' \in T(\Sigma)$, $t\ \mathsf{R}\ t'$ implies $t' \in Q$ (i.e. $\mathsf{R}(Q) \subseteq Q$). If a set Q is R-closed we write R-closed(Q). It is easy to verify that if two relation $\mathsf{R},\mathsf{R}' \subseteq T(\Sigma)\times T(\Sigma)$ are such that $\mathsf{R}' \subseteq \mathsf{R}$, then for all set $Q \subseteq T(\Sigma)$, R-closed(Q) implies R'-closed(Q).

Definition 9. *A relation $\mathsf{R} \subseteq T(\Sigma)\times T(\Sigma)$ is a bisimulation if R is symmetric and for all $t,t' \in T(\Sigma)$, $\pi \in \Delta(T(\Sigma))$, $a \in A$,*

$$t\ \mathsf{R}\ t' \text{ and } t \xrightarrow{a} \pi \text{ imply that there exists } \pi' \in \Delta(T(\Sigma)) \text{ s.t. } t' \xrightarrow{a} \pi' \text{ and } \pi\ \mathsf{R}\ \pi',$$

where $\pi\ \mathsf{R}\ \pi'$ if and only if $\forall Q \subseteq T(\Sigma) : \mathsf{R}\text{-closed}(Q) \Rightarrow \pi(Q) = \pi'(Q)$. We define bisimilarity $\sim$ as the smallest relation that includes all other bisimulations. It is well-known that $\sim$ is itself a bisimulation and an equivalence relation.

Let $\{Y_l\}_{l\in L}$ be a family of sets of term variables with the same cardinality. The l-th element of a tuple $\vec{y}$ is denoted by $\vec{y}(l)$. For a set of tuples $T = \{\vec{y_i} \mid i \in I\}$ we denote the l-th projection by $\pi_l(T) = \{\vec{y_i}(l) \mid i \in I\}$. Fix a set $\mathsf{Diag}\{Y_l\}_{l\in L} \subseteq \prod_{l\in L} Y_l$ such that:

(i) for all $l \in L$, $\pi_l(\mathsf{Diag}\{Y_l\}_{l\in L}) = Y_l$; and

(ii) for all $\vec{y},\vec{y'} \in \mathsf{Diag}\{Y_l\}_{l\in L}$, $(\exists l \in L : \vec{y}(l) = \vec{y'}(l)) \Rightarrow \vec{y} = \vec{y'}$.

Property (ii) ensures that different $\vec{y},\vec{y'} \in \mathsf{Diag}\{Y_l\}_{l\in L}$ differ in all positions and by property (i) every variable of every Y_l is used in one $\vec{y} \in \mathsf{Diag}\{Y_l\}_{l\in L}$. Diag stands for "diagonal", following the intuition that each $\vec{y}$ represents a coordinate in the space $\prod_{l\in L} Y_l$, then $\mathsf{Diag}\{Y_l\}_{l\in L}$ can be seen as the line that traverses the main diagonal of the space. Notice that, letting L be a natural number, for $Y_l = \{y_l^0, y_l^1, y_l^2, \ldots\}$ a possible definition for $\mathsf{Diag}\{Y_l\}_{l\in L}$ is $\mathsf{Diag}\{Y_l\}_{l\in L} = \{(y_0^0, y_1^0, \ldots, y_L^0), (y_0^1, y_1^1, \ldots, y_L^1), (y_0^2, y_1^2, \ldots, y_L^2), \ldots\}$.

Definition 10. *Let $P = (\Sigma, A, R)$ be a PTSS. A rule $r \in R$ is in ntμfθ* format *if it has the following form*

$$\frac{\bigcup_{m\in M}\{t_m(\vec{z}) \xrightarrow{a_m} \mu_m^{\vec{z}} : \vec{z} \in \mathcal{Z}\} \cup \bigcup_{n\in N}\{t_n(\vec{z}) \xrightarrow{b_n} \not\to : \vec{z} \in \mathcal{Z}\} \cup \{\theta_l(Y_l) \trianglerighteq_{l,k} p_{l,k} : l \in L, k \in K_l\}}{f(x_1,\ldots,x_{\mathrm{r}(f)}) \xrightarrow{a} \theta}$$

with $\trianglerighteq_{l,k} \in \{>,\geq\}$ for all $l \in L$ and $k \in K_l$, and it satisfies the following conditions:

1. *Each set Y_l should be at least countably infinite, for all $l \in L$, and the cardinality of L should be strictly smaller than that of the Y_l's.*
2. *$\mathcal{Z} = \mathsf{Diag}\{Y_l\}_{l\in L} \times \prod_{w\in W}\{w\}$, with $W \subseteq \mathcal{V} \setminus \bigcup_{l\in L} Y_l$.*
3. *All variables $\mu_m^{\vec{z}}$, with $m \in M$ and $\vec{z} \in \mathcal{Z}$, are different.*
4. *For all $\vec{z},\vec{z'} \in \mathcal{Z}$, $m \in M$, if $\mu_m^{\vec{z}},\mu_m^{\vec{z'}} \in Var(\theta) \cup (\cup_{l\in L} Var(\theta_l))$ then $\vec{z} = \vec{z'}$.*
5. *For all $l \in L$, $Y_l \cap \{x_1,\ldots,x_{\mathrm{r}(f)}\} = \emptyset$, and $Y_l \cap Y_{l'} = \emptyset$ for all $l' \in L$, $l \neq l'$.*
6. *All variables $x_1,\ldots,x_{\mathrm{r}(f)}$ are different.*
7. *For all $l \in L$, $Var(\theta_l) \cap (\{x_1,\ldots,x_{\mathrm{r}(f)}\} \cup \bigcup_{l'\in L} Y_{l'}) = \emptyset$.*
8. *$f \in F$ and for all $m \in M$ and $n \in N$, $t_m, t_n \in \mathbb{T}(\Sigma)$. In all cases, if $t \in \mathbb{T}(\Sigma)$ and $Var(t) \subseteq \{w_1,\ldots,w_H\}$, $t(w'_1,\ldots,w'_H)$ is the same term as t where each occurrence of variable w_h (if it appears in t) has been replaced by variable w'_h, for $1 \leq h \leq H$.*

9. $\theta, \theta_l \in \mathbb{DT}(\Sigma)$ *for all* $l \in L$.

A rule $r \in R$ *is in* $nt\mu x\theta$ format *if its form is like above but has a conclusion of the form* $x \xrightarrow{a} \theta$ *and, in addition, it satisfies the same conditions as above only that whenever we write* $\{x_1, \ldots, x_{\mathrm{r}(f)}\}$*, we should write* $\{x\}$*. A rule* $r \in R$ *is in* $nx\mu f\theta$ format *if it is in* $nt\mu f\theta$ *format and the sources of its positive premises are term variables. P is in* $nt\mu f\theta$ *(resp.* $nt\mu x\theta$*,* $nx\mu f\theta$*)* format *if all its rules are in* $nt\mu f\theta$ *(resp.* $nt\mu x\theta$*,* $nx\mu f\theta$*) format. P is in* $nt\mu f\theta/nt\mu x\theta$ format *if each of its rules is either in* $nt\mu f\theta$ *format or* $nt\mu x\theta$ *format.*

The rationale behind each of the restrictions are discussed in [7] in depth. In the following we briefly summarize it. Term variables $x_1, \ldots, x_{\mathrm{r}(f)}$ appearing in the source of the conclusion are binding. Variables in $\bigcup_{l \in L} Y_l$ and those appearing in instantiable Dirac distributions are also binding when appearing in quantitative premises. Therefore they need to be all different. This is stated in conditions 3, 5, and 7. Distribution variables in $\{\mu_m^{\vec{z}} \mid m \in M \wedge \vec{z} \in \mathcal{Z}\}$ are also binding when appearing on the target of a positive premise. Hence they also need to be different, which is stated in condition 6. If Y_l is finite, quantitative premises will allow to count the minimum number of terms that gather certain probabilities. This goes against the spirit of bisimulation that measures equivalence classes of terms regardless of the size of them. Therefore Y_l needs to be infinite (condition 1). Condition 4 is more subtle; together with each set of premises $\{t_m(\vec{z}) \xrightarrow{a_m} \mu_m^{\vec{z}} : \vec{z} \in \mathcal{Z}\}$ it ensures a symmetric behaviour of terms $t_m(\vec{z})$ for every possible instantiation of variables $\vec{z}$. A clear example that shows the need for this symmetry is provided in [7]. The need for the source of the conclusion and targets of positive premises to have a particular shape is the same as in the *tyft/tyxt* format [14]. Conditions 2, 8, and 9 are actually notations and definitions.

The definition provided here corrects some mistakes inadvertently introduced in the $nt\mu fv/nt\mu xv$ format in [7], more precisely on the quantitative premises and condition 4 in Def. 11 (which corresponds to our condition 4). Another mistake in [7] was omitting to require that PTSS are well-founded as hypothesis for the congruence theorem. This is corrected in the following, where we extend the congruence theorem to the $nt\mu f\theta/nt\mu x\theta$ format. [1]

Definition 11. *Let W be a set of positive and quantitative premises. The* dependency directed graph *of W is given by* $G_W = (V, E)$ *with* $V = \cup_{\psi \in W} Var(\psi)$ *and* $E = \{\langle x, \mu\rangle \mid t \xrightarrow{a} \mu, x \in Var(t)\} \cup \{\langle \zeta, y\rangle \mid (\theta(Y) \trianglerighteq p) \in W, \zeta \in Var(\theta), y \in Y\}$*. We say that W is* well-founded *if any backward chain of edges in* G_W *is finite. Define for each* $x \in V$*,* $n_{\mathrm{VDG}}(x) = \sup(\{n_{\mathrm{VDG}}(y) + 1 \mid (y, x) \in E\})$*, where* $\sup(\emptyset) = 0$*. A rule is called* well-founded *if its set of positive and quantitative premises is well-founded. A PTSS is called* well-founded *if all its rules are well-founded.*

Theorem 4. *Let P be a well-founded stratifiable PTSS in* $nt\mu f\theta/nt\mu x\theta$ *format. Then* $\sim$ *is a congruence relation for all operators defined in P.*

5 $nt\mu f\theta/nt\mu x\theta$ format reduces to pntree

The reduction procedure requires results from unification theory over infinite domains. Instead using the result presented in [8], we use the variation presented in [9, Lemma 3.2] that proves some extra properties needed to prove our main result.

Definition 12. *A substitution* σ *is a* unifier *for a substitution* ρ *if* $\sigma\rho = \sigma$*. In this case, we say that* ρ *is* unifiable.

[1] Both issues are explained in detail in the corrigendum of [7]: http://cs.famaf.unc.edu.ar/ lee/publications/corrigendum-Fossacs2012.pdf

Lemma 6. *If a substitution ρ is unifiable, then there is a unifier $\hat{\sigma}$ for ρ such that: (i) each unifier σ for ρ is also a unifier for $\hat{\sigma}$ (ii) if $\rho(\zeta) = \zeta$ then $\hat{\sigma}(\zeta) = \zeta$, for all $\zeta \in \mathcal{V} \cup \mathcal{M}$, and (iii) if $\rho^n(\zeta)$ is a variable for all $n \geq 0$ then $\hat{\sigma}(\zeta)$ is a variable. We call $\hat{\sigma}$ the most general unifier.*

The main theorem 5 showing that every PTSS in *ntμfθ/ntμxθ*-format can be reduced to a transition equivalent PTSS in pntree format is developed incrementally. First of all, we show that every *ntμxθ*-rule can be expressed by a set of *ntμfθ*-rules by replacing the source variable of the conclusion with an appropriate context $f(\vec{x})$ (Lemma 7). Secondly, we show that for all PTSS P in *ntμfθ* format there is a PTSS P' in *nxμfθ* format such that $P \vdash \frac{H}{c}$ iff $P' \vdash \frac{H}{c}$ for all rules $\frac{H}{c}$ in *nxμfθ* format (Lemma 8). Notice that this result implies that $P \vdash \frac{H'}{c}$ iff $P' \vdash \frac{H'}{c}$ for all rule $\frac{H'}{c}$ with H' a set of closed negative premises, then by Lemma 5, P and P' are equivalent. Finally, we prove that for all PTSS P in *nxμfθ* format there is a PTSS P' in pntree format (a PTSS in well-founded *nxμfθ* format without free variables), such that for every closed transition rule $\frac{H}{c}$ with only negative premises, $P \vdash \frac{H}{c}$ iff $P' \vdash \frac{H}{c}$ (Lemma 9). Again, by Lemma 5, P and P' are equivalent. This series of lemmas leads to the main theorem stating that every PTSS consisting of rules in the *ntμfθ/ntμxθ* format can be reduced to a transition equivalent PTSS in the more restrictive pntree format. Furthermore, this shows also that the rules of a PTSS in *ntμfθ/ntμxθ* format do not have to be well-founded in order to guarantee that the bisimilarity of the induced PTS is a congruence.

The reduction of proof structures follows the logic of [9]. In the probabilistic setting we need to treat additionally quantitative premises as follows: While substitutions replace distribution variables by distribution terms the substitution $\rho(\theta(Y) > p)$ leads to a well-defined quantitative literal (ρ is defined as $\rho(y) = y$ for all $y \in Y$). Because by construction σ unifies ρ we have that whenever $\sigma(\theta(Y) > p)$ then also $\sigma(\rho(\theta(Y) > p))$. This shows the satisfaction of the quantitative premises.

Lemma 7. *Let $P = (\Sigma, A, R)$ be a stratifiable PTSS in ntμfθ/ntμxθ format. Then there is a stratifiable PTSS $P' = (\Sigma, A, R')$ in ntμfθ format that is transition equivalent to P.*

Lemma 8. *Let $P = (\Sigma, A, R)$ be a PTSS in ntμfθ format. Then there is a PTSS $P' = (\Sigma, A, R')$ in nxμfθ format such that $P \vdash \frac{H}{c}$ iff $P' \vdash \frac{H}{c}$ for all rule $\frac{H}{c}$ in nxμtθ format. (A rule is in nxμtθ format if the source of every positive premise is a term variable and its target is a distribution variable.)*

Proof. Define $P' = (\Sigma, A, R')$ such that $r \in R'$ iff r is a provable rule from P in *nxμfθ* format. The right to left implication follows straightforward from Lemma 2.

For the left to right implication we proceed by induction on the partial order over proof structures. Suppose $P \vdash \frac{H}{c}$, with a rule $\frac{H}{c}$ in *nxμtθ* format, and let (B, r, ϕ) be a proof structure for $\frac{H}{c}$ over P. Then by Def. 6 there is substitution σ s.t. (a) $\sigma(\mathrm{top}(B, r, \phi) - \mathrm{qtop}(B, r, \phi)) \subseteq H$, (b) closed quantitative premise in $\sigma(\mathrm{qtop}(B, r, \phi))$ hold, (c) open quantitative premise in $\sigma(\mathrm{qtop}(B, r, \phi))$ belong to H, and (d) $\sigma(\mathrm{conc}(r)) = c$.

From (B, r, ϕ) we construct recursively a substructure (B', r, ϕ') which is a proof structure for a rule $r' \in R'$, i.e. r' is in *nxμfθ* format, such that $\sigma(\mathrm{conc}(r')) = c$ and for each premise c' of $\sigma(r')$ the rule $\frac{H}{c'}$ is provable from R' i.e. $\frac{H}{c} \in R'^{\vdash}$ or c' is a valid closed quantitative literal. Then, by Lemma 1, $\frac{H}{c}$ is provable from P'. Furthermore, we construct a partial substitution ρ which is unified by σ, i.e. if $\rho(x)$ is defined then $\sigma(\rho(x)) = \sigma(x)$. In this construction ρ^0 is defined as the identity function. We proceed with the definitions of the transition rules B' and the substitution ρ:

(i) $r \in B'$.

(ii) If $b \in B \setminus \{r\}$ and $\phi(b)$ is a premise $t_m(\vec{z}) \xrightarrow{a_m} \mu_m^{\vec{z}}$ of a rule in B' s.t there is $k \geq 0$ with:
   (a) $\rho^i(t_m(\vec{z}))$ is defined for $i = 0, \ldots, k$
   (b) $\rho^i(t_m(\vec{z}))$ are variables for $i = 0, \ldots, k-1$

(c) $\rho^k(t_m(\vec{z}))$ has the form $f(t_1,\dots,t_{\mathsf{r}(f)})$ with $t_i \in \mathbb{T}(\Sigma)$

then $b \in B'$. Notice that the conditions can be satisfied only if $\rho^i(t_m(\vec{z}))$ is a variable for $i = 0,\dots,k-1$. Moreover $\rho^0(t_m(\vec{z})) = t_m(\vec{z})$ is a variable. In addition, this variable belongs to $\vec{z}$.

(iii) Since σ matches with (B,r,ϕ), $\sigma(\mathrm{conc}(b)) = \sigma(t_m(\vec{z}) \xrightarrow{a_m} \mu_m^{\vec{z}})$. Because the rule format restricts the form of the conclusion $\mathrm{conc}(b)$, then we can rewrite the last equality by: $\sigma(f(x_1,\dots,x_{\mathsf{r}(f)}) \xrightarrow{a} \theta) = \sigma(t_m(\vec{z}) \xrightarrow{a_m} \mu_m^{\vec{z}})$ In addition, σ unifies the partial substitution ρ, then if $\rho^{k-1}(t_m(\vec{z}))$ is a variable it holds: $\sigma(t_m(\vec{z})) = \sigma\rho^k(t_m(\vec{z})) = \sigma(f(t_1,\dots,t_n))$.

Because $\mathrm{conc}(b)$ has the form $f(x_1,\dots,x_{\mathsf{r}(f)}) \xrightarrow{a} \theta$ it holds $\sigma(x_j) = \sigma(t_j)$ for $j = 1,\dots,\mathsf{r}(f)$ and $\sigma(\theta) = \sigma(\mu_m^{\vec{z}})$. Define $\rho(x_j) = t_j$ for $j = 1,\dots,\mathsf{r}(f)$ (here we define the left side of a conclusion of a rule in $B' \setminus r$). Besides, define $\rho(\mu_m^{\vec{z}}) = \theta$. Notice that this extension of ρ is unified by σ and, by Def. 6, the variables x_j and $\mu_m^{\vec{z}}$ appear only in this rule, then we are not redefining substitution ρ.

(iv) Define $\rho(\zeta) = \zeta$ for all variable ζ if ζ is not defined for ρ. Substitution σ unifies this extension of ρ.

(v) Finally, ϕ' is the restriction of ϕ to $B' \setminus \{r\}$. (Notice that the substitution ρ is defined for the the right side of a positive premise in the image of ϕ' in item (iii).)

Substitution σ unifies substitution ρ, by Lemma 6, there is a substitution ρ' which unifies ρ and:

(ρ'i) $\sigma\rho' = \sigma$.

(ρ'ii) If $\rho(\zeta) = \zeta$ then $\rho'(\zeta) = \zeta$, with ζ a term or distribution variable.

(ρ'iii) If $\rho^k(\zeta)$ is a variable for $k \geq 0$ then $\rho'(\zeta)$ is a variable.

The proof structure (B',r,ϕ') and the substitution ρ' are completely defined, now we can prove that ρ' matches with (B',r,ϕ'). Let b a rule used to construct B' and consider the substitution ρ. Recall that the conclusion of b has the form $f(x_1,\dots,x_{\mathsf{r}(f)}) \xrightarrow{a} \theta$ and $\phi'(b) = t_m(\vec{z}) \xrightarrow{a_m} \mu_m^{\vec{z}}$ is such that $\rho^k(t_m(\vec{z})) = f(t_1,\dots,t_{\mathsf{r}(f)}) = \rho(f(x_1,\dots,x_{\mathsf{r}(f)}))$ by (ii) and the definition of ρ for x_i in (iii). Since ρ' unifies ρ then

$$\begin{array}{rcccccl} \rho'(\phi'(b)) & = & \rho'(t_m(\vec{z}) \xrightarrow{a_m} \mu_m^{\vec{z}}) & = & \rho'(\rho^k(t_m(\vec{z})) \xrightarrow{a_m} \rho(\mu_m^{\vec{z}}))) & = & \\ & = & \rho'(\rho(f(x_1,\dots,x_{\mathsf{r}(f)}) \xrightarrow{a_m} \theta) & = & \rho'(f(x_1,\dots,x_{\mathsf{r}(f)}) \xrightarrow{a_m} \theta) & = & \rho'(\mathrm{conc}(b)) \end{array}$$

Then the substitution ρ' matches with the proof structure (B',r,ϕ').

To show that the rule $s = \rho'\left(\dfrac{\{h : h \in \mathrm{top}(B',r,\phi'), h \text{ is not a closed quantitative premise}\}}{\mathrm{conc}(r)}\right)$ is provable (Def. 6), it remains to show that if a quantitative premise in $\mathrm{qtop}(B',r,\phi')$ is closed then it is also valid. Let $\psi \in \mathrm{qtop}(B',r,\phi')$ be a quantitative premise. Then if $\rho'(\psi)$ is closed, since σ unifies ρ', it holds that $\sigma(\psi) = \sigma(\rho'(\psi)) = \rho'(\psi)$, which implies that also $\sigma(\psi)$ is a closed literal. Because the rule $\sigma\left(\dfrac{\{h : h \in \mathrm{top}(B,r,\phi), h \text{ is not a closed quantitative premise}\}}{\mathrm{conc}(r)}\right)$ is provable we have that $\sigma(\psi)$ holds and therefore also $\rho'(\psi)$ holds.

Finally we prove that the rule s is in $nx\mu f\theta$ format. From the construction by ρ we know that if x is s.t. $\rho(x) \neq x$ then x satisfies one of the following conditions:

1. x appears in the left-hand side of a conclusion of a rule in $B' \setminus \{r\}$,
2. x appears in the right-hand side of a positive premise in the image of ϕ'.

Then if $g(x_1,\dots,x_m) \xrightarrow{b} \theta$ is the conclusion of r, $\rho(x_j) = x_j$ for $j = 1,\dots,m$ and, hence $\rho'(x_j) = x_j$ because of (ρ'ii). On the other hand, if $\zeta \in Var(\theta)$ is a variable that appears in the right-hand side of

a positive premise in the image of ϕ', i.e. ζ is a distribution variable, we have $\rho'(\zeta) \in \mathbb{DT}(\Sigma)$ and then $\rho'(\zeta) \in \mathbb{DT}(\Sigma)$. Therefore the conclusion $\rho'(g(x_1,\ldots,x_m) \xrightarrow{b} \theta)$ of s has the form $g(x_1,\ldots,x_m) \xrightarrow{a} \rho'(\theta)$ as the $nx\mu f\theta$ format demands.

We continue with the premises of s. Let $\rho'(t \xrightarrow{a} \mu)$ be a positive premise in $\rho'(\mathrm{top}(B', r, \phi'))$ then $t \xrightarrow{a} \mu$ is a positive premise of a rule in B' which does not belong to the image of ϕ'. Then μ is such that $\rho(\mu) = \mu$ and this implies $\rho'(\mu) = \mu$. To prove that $\rho'(t)$ is a variable there are 2 cases to investigate:

- $t \xrightarrow{a} \mu \in \mathrm{top}(B, r, \phi)$. Then $\sigma(t \xrightarrow{a} \mu) \in H$ and because $\frac{H}{c}$ is in $nx\mu t\theta$ format, then $\sigma(t)$ is a variable. Therefore $\sigma\rho'(t) = \sigma(t)$ and then $\rho'(t)$ is a variable.
- $t \xrightarrow{a} \mu \notin \mathrm{top}(B, r, \phi)$. Then there is a rule b s.t. $\phi(b) = t \xrightarrow{a} \mu$. Since $t \xrightarrow{a} \mu$ does not belong to the image of ϕ' we have that $b \notin B'$. By B' and the construction of ρ we have that $\rho^k(t)$ is a variable for all $k \geq 0$. Then (ρ'iii) ensures that $\rho'(t)$ is a variable.

This shows that the positive premises also fulfill the requirements of the $nx\mu f\theta$ format.

We proceed with the quantitative premises. Let $(\theta(Y) \trianglerighteq p) \in \mathrm{qtop}(B', r, \phi')$ with $\theta \in \mathbb{DT}(\Sigma)$. By the same reasoning as applied for the target of the conclusion we get $\rho'(\theta) \in \mathbb{DT}(\Sigma)$. In addition, $\rho(y) = y$ for all $y \in Y$ because they do not appear in the left-hand side of a conclusion, and hence $\rho'(y) = y$. Thus, $\rho'(\theta(Y) \trianglerighteq p)$ has the proper form.

Syntactical restriction for positive and quantitative premises and conclusion are satisfied. Besides, there is no restriction for negative premises, therefore s is in $nx\mu f\theta$ format and then $s \in R'$.

For all positive premises $c' \in \sigma(\mathrm{top}(B', r, \phi'))$ the rule $\frac{H}{c'}$ is in $nx\mu t\theta$ and it is provable in R by a proof sub-structure smaller than (B, r, ϕ). Thus, by induction we get that these rules are provable in R'. Applying Lemma 1 on these rules and s shows that $\frac{H}{c}$ is provable in R'. □

Definition 13. *We say that a variable x occurs* free *in a rule r if it occurs in r but not in the source of the conclusion nor in W_j with $\theta_j(W_j) \trianglerighteq_j q_j \in qprem(r)$. We say that a distribution variable μ occurs* free *in a rule r if it occurs in r but not in the target of a positive premise.*

Definition 14. *A PTSS $P = (\Sigma, A, R)$ is in* pntree format *if all rules in R are well-founded $nx\mu f\theta$ rules without free variables.*

Lemma 9. *Let $P = (\Sigma, A, R)$ be a PTSS in $nx\mu f\theta$ format. Then there is a PTSS $P' = (\Sigma, A, R')$ in pntree format such that for every closed transition rule $\frac{H}{c}$ with only negative premises, $P \vdash \frac{H}{c}$ iff $P' \vdash \frac{H}{c}$*

Proof. Let $P' = (\Sigma, A, R')$ such that R' is the set of provable rules from P in pntree format. By Lemma 2, the right to left implication holds.

For the left to right implication we proceed by induction. Let $\frac{H}{c}$ be closed with H containing negative literals only. Let $\frac{H}{c}$ be provable from P, i.e. $\frac{H}{c} \in R^{\vdash}$. Then either $c \in H$, c is a valid closed quantitative literal, or there is a rule r and a substitution ρ such that $\rho(\mathrm{conc}(r)) = c$ and, for all premises $c' \in \rho(\mathrm{pprem}(r))$, $\frac{H}{c'} \in R^{\vdash}$. Then $\frac{H}{c'} \in R'^{\vdash}$ either trivially or by induction.

Because r is $nx\mu f\theta$ format, r has the form

$$\frac{\bigcup_{m\in M}\{w_m^{\vec{z}} \xrightarrow{a_m} \mu_m^{\vec{z}} : \vec{z} \in \mathcal{Z}\} \cup \bigcup_{n\in N}\{t_n(\vec{z}) \xrightarrow{b_n} \not\rightarrow : \vec{z} \in \mathcal{Z}\} \cup \{\theta_l(Y_l) \trianglerighteq_{l,k} p_{l,k} : l \in L, k \in K_l\}}{f(x_1,\ldots,x_{\mathrm{r}(f)}) \xrightarrow{a} \theta}$$

where each $w_m^{\vec{z}}$ is a variable in $\vec{z}$.

Let G be the variable dependency graph associated to $\mathrm{pprem}(r) \cup \mathrm{qprem}(r)$. From r, we construct a rule $r' \in S$ as follows. Let $\mu_m^{\vec{z}}$ be the target of a positive premise such that there is no backward path in

G from a vertex $\mu^{\vec{z}}_m$ to some vertex x_i, with $i \in \{1,\dots,x_{\mathrm{r}(f)}\}$. Notice that, by the symmetry requirements in Def. 10, this happens for all $\mu^{\vec{z}}_m$ with $\vec{z} \in \mathcal{Z}$. We first obtain a rule r'' by (i) replacing variables $w^{\vec{z}}_m$ and $\mu^{\vec{z}}_m$ by $\rho(w^{\vec{z}}_m)$ and $\rho(\mu^{\vec{z}}_m)$, respectively, and (ii) replacing every free variable ζ in θ_l and θ by $\rho(\zeta)$. The resulting rule r'' does not have free variables and it is a substitution instance of r, so r'' is provable from P. To obtain r', replace each closed positive premise $\rho(w^{\vec{z}}_m \xrightarrow{a} \mu^{\vec{z}}_m)$ by H. Since, $w^{\vec{z}}_m \xrightarrow{a} \mu^{\vec{z}}_m$ is a positive premise of r, $\frac{H}{\rho(w^{\vec{z}}_m \xrightarrow{a} \mu^{\vec{z}}_m)} \in R^{\vdash}$. Then r' is also provable from P.

Notice that the resulting rule r' is in $nx\mu f\theta$ format without free variables. Morever, r' is well-founded since any dependency backward chain ends in a vertex x_i. Hence r' is a pntree rule and therefore $r' \in R'$.

Let $p \in \mathrm{prem}(r')$. Then either $p \in H$ (and hence p is closed) or $p \in \mathrm{prem}(r)$. In any case, $\frac{H}{\rho(p)} \in R'^{\vdash}$ (if $p \in \mathrm{prem}(r)$, it follows by induction). Therefore $\frac{H}{\rho(\mathrm{conc}(r'))} \in R'^{\vdash}$. Since $\rho(\mathrm{conc}(r')) = \rho(\mathrm{conc}(r)) = c$, $\frac{H}{c} \in R'^{\vdash}$. □

Theorem 5. *Let $P = (\Sigma, A, R)$ be a PTSS in $nt\mu f\theta/nt\mu x\theta$ format. There is a PTSS $P' = (\Sigma, A, R')$ in pntree format that is transition equivalent to P.*

The proof of Theorem 5 follows by applying Lemmas 7, 8, 9, and 5, in that order.

Let P be a stratifiable PTSS in $nt\mu f\theta/nt\mu x\theta$ format and let S be its stratification. If r is a provable rule from P, conditions (i) and (ii) in Def. 3 also hold for stratification S in rule r. (This can be shown by induction.) Then, S is also a stratification for the PTSS P' in pntree format obtained as in Theorem 5. Since pntree rules are well-founded $nt\mu f\theta$ rules, from Theorems 4 and 5, we have the following corollary.

Corollary 1. *If P is a stratifiable PTSS in $nt\mu f\theta/nt\mu x\theta$ format, $\sim$ is a congruence for all operators in P.*

To conclude the section, we remark that negative premises cannot be reduced to variables. Following the nomenclature of [9], we say that a rule is in *simple pntree format* if it is in pntree format and all its negative premises have the form $x \overset{a}{\not\to}$. It turns out that the pntree format (and hence also the $nt\mu f\theta/nt\mu x\theta$ format) is strictly more expressive than simple pntree format. We will not dwell on this since example and rationale of the difference of expressiveness in the non-probabilistic case applies mutatis mutandi to our case (see [9]).

6 Concluding remarks

We introduced the rule format $nt\mu f\theta/nt\mu x\theta$ which enriches $nt\mu fv/nt\mu xv$ [7] by allowing distribution terms to appear in quantitative premises and conclusions of rules. We showed that it ensures that bisimulation equivalence is a congruence for operators of well-founded PTSSs. On proving this, we corrected a mistake introduced in [7]. The richer syntactic structure of the quantitative premises and the conclusion of the rules allows us to define a reduction of $nt\mu f\theta/nt\mu x\theta$ PTSSs to a transition equivalent PTSS consisting of only pntree rules. This construction confirms that the well-foundedness requirement in $nt\mu f\theta/nt\mu x\theta$ is not necessary to guarantee that bisimilarity is a congruence.

We already know that the $nt\mu f\theta/nt\mu x\theta$ format is equally expressive if restricted to quantitative premises of the form $\theta(Y) > q$ with $q \in [0,1] \cap \mathbb{Q}$. However, we do not know whether distribution terms are really needed. We actually suspect that they are, and hence, that the $nt\mu f\theta/nt\mu x\theta$ format is strictly more expressive than the $nt\mu fv/nt\mu xv$ format.

Pntree rules are nearly ruloids [5] except that negative premises may still contain non-variable terms. The decomposition method of [4, 10] to develop modular compositional proof systems can be adapted to pntree rules by applying the negation-as-failure semantics for the logical characterization of negative premises of pntree rules. This will allow us to derive expressive congruence formats for probabilistic

behavioral equivalences from their logical characterization in a structured way, following the approach of [4].

Both [7] and this work have opened a new way of thinking about probabilistic transition system specifications. One of the nicest things is that the *ntμfθ/ntμxθ* follows quite closely the structure of non-probabilistic formats (particularly, *ntyft/ntyxt*). Hence, many ideas for further work can be borrowed from the non-probabilistic setting.

References

[1] Jos C. M. Baeten, Jan A. Bergstra & Scott A. Smolka (1995): *Axiomatizing Probabilistic Processes: ACP with Generative Probabilities*. Inf. Comput. 121(2), pp. 234–255, doi:10.1006/inco.1995.1135.

[2] Falk Bartels (2002): *GSOS for Probabilistic Transition Systems*. Electr. Notes Theor. Comput. Sci. 65(1).

[3] Falk Bartels (2004): *On Generalised Coinduction and Probabilistic Specification Formats*. Ph.D. thesis, Vrije Universiteit.

[4] Bard Bloom, Wan Fokkink & Rob van Glabbeek (2004): *Precongruence formats for decorated trace semantics*. ACM TOCL 5, pp. 26–78, doi:10.1145/963927.963929.

[5] Bard Bloom, Sorin Istrail & Albert R. Meyer (1995): *Bisimulation Can't be Traced*. J. ACM 42(1), pp. 232–268, doi:10.1145/200836.200876.

[6] Roland Bol & Jan Friso Groote (1996): *The meaning of negative premises in transition system specifications*. J. ACM 43(5), pp. 863–914, doi:10.1145/234752.234756.

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

[8] Wan Fokkink (1997): *Unification for infinite sets of equations between finite terms*. Information Processing Letters 62(4), pp. 183 – 188, doi:10.1016/S0020-0190(97)00063-X.

[9] Wan Fokkink & Rob J. van Glabbeek (1996): *Ntyft/Ntyxt Rules Reduce to Ntree Rules*. Inf. Comput. 126(1), pp. 1–10, doi:10.1006/inco.1996.0030.

[10] Daniel Gebler & Wan Fokkink (2012): *Compositionality of Probabilistic Hennessy-Milner Logic through Structural Operational Semantics*. In: Proc. CONCUR 2012, LNCS 7454, Springer, pp. 395–409.

[11] 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.

[12] Rob J. van Glabbeek, Scott A. Smolka & Bernhard Steffen (1995): *Reactive, Generative and Stratified Models of Probabilistic Processes*. Inf. Comput. 121(1), pp. 59–80, doi:10.1006/inco.1995.1123.

[13] 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.

[14] Jan Friso Groote & Frits Vaandrager (1992): *Structured operational semantics and bisimulation as a congruence*. Inf. Comput. 100(2), pp. 202–260, doi:10.1016/0890-5401(92)90013-6.

[15] Bartek Klin & Vladimiro Sassone (2008): *Structural operational semantics for stochastic process calculi*. In: FoSSaCS, LNCS 4962, Springer, pp. 428–442, doi:10.1007/978-3-540-78499-9_30.

[16] Ruggero Lanotte & Simone Tini (2009): *Probabilistic bisimulation as a congruence*. ACM Trans. Comput. Log. 10(2), doi:10.1145/1462179.1462181.

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

[18] Mohammad Reza Mousavi, Michel A. Reniers & Jan Friso Groote (2007): *SOS formats and meta-theory: 20 years after*. Theor. Comput. Sci. 373(3), pp. 238–272, doi:10.1016/j.tcs.2006.12.019.

[19] Gordon D. Plotkin (1981): *A structural approach to operational semantics*. Report DAIMI FN-19, Aarhus University, doi:10.1016/j.jlap.2004.05.001. Reprinted in *J. Log. Algebr. Program.*, 60-61:17-139, 2004.

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