

# Implementing the Stochastics Brane Calculus in a Generic Stochastic Abstract Machine

Marino Miculan      Iliaria Sambarino

Department of Mathematics and Computer Science, University of Udine, Italy

marino.miculan@uniud.it

ilaria.sambarino@gmail.com

In this paper, we deal with the problem of implementing an abstract machine for a stochastic version of the Brane Calculus. Instead of defining an *ad hoc* abstract machine, we consider the generic stochastic abstract machine introduced by Lakin, Paulevé and Phillips. The nested structure of membranes is flattened into a set of species where the hierarchical structure is represented by means of *names*. In order to reduce the overhead introduced by this encoding, we modify the machine by adding a *copy-on-write* optimization strategy. We prove that this implementation is adequate with respect to the stochastic structural operational semantics recently given for the Brane Calculus. These techniques can be ported also to other stochastic calculi dealing with nested structures.

## 1 Introduction

A fundamental issue in Systems Biology is modelling the membrane interaction machinery. Several models have been proposed in the literature [11, 14, 3]; among them, the Brane Calculus [4] has been arisen as a good model focusing on abstract membrane interactions, still being sound with respect to biological constraints (e.g. bitonality). In this calculus, a process represents a system of nested compartments, where active components are *on* membranes, not inside them. This reflects the biological evidence that functional molecules (proteins) are embedded in membranes, with consistent orientation.

In the original definition of the Brane Calculus [4] (which we will recall in Section 2) membranes interact according to three basic reaction rules corresponding to *phagocytosis*, *endo/exocytosis*, and *pinocytosis*. However, this semantics does not take into account quantitative aspects, like stochastic distributions, which are important for, e.g., implementing stochastic simulations.

A stochastic semantics for the Brane Calculus has been provided in [2], following an approach pioneered in [5] (but see also [8, 12] for Markov processes). Instead of giving a stochastic version  $P \xrightarrow{a,r} Q$  of the reaction relation, in this semantics each process is given a *measure* of the stochastic distribution of the possible outcomes. More precisely, we define a relation  $P \rightarrow \mu$  associating to a process  $P$  an action-indexed family of measures  $\mu$ : for an action  $a$ , the measure  $\mu_a$  specifies for each measurable set  $S$  of processes, the rate  $\mu_a(S) \in \mathbb{R}^+$  of  $a$ -transitions from  $P$  to (elements of)  $S$ . An advantage of this approach is that we can apply results from measure theory for solving otherwise difficult issues, like instance-counting problems; moreover, process measures are defined *compositionally*, and in fact the relation  $P \rightarrow \mu$  can be characterized by means of a set of rules in a GSOS-like format. We will recall this stochastic semantics and its main properties in Section 3.

In this paper, we use this new semantics for defining a *stochastic abstract machine* for the Brane Calculus, so that it can be effectively used for *in silico* simulations of membrane systems. Defining an *ad hoc* abstract machine for the Brane Calculus would be a complex task; instead, we take advantage of the *generic abstract machine for stochastic process calculi* (GSAM for short) introduced in [13, 10] as a general tool for simulating a broad range of calculi. This machine can be instantiated to a particular

calculus by defining a function for transforming a process of the calculus to a set of *species*, and another for computing the set of possible reactions between species.

An important aspect is that this abstract machine does not have a native notion of compartment, which is central in the Brane Calculus (as in any other model of membranes). To overcome this problem, we adopt a “flat” representation of membrane systems, used also in [10], where the hierarchical structure is represented by means of *names*: each name represents a compartment, and each species is labelled with the name of the compartment where it is located, and the name of its inner compartment (if any). So names and species are the nodes and the arcs of the tree, respectively. This technique can be used for representing any system with a tree-like structure of compartments.

However, this approach does not scale well, as the population of species may grow enormously: for instance, a population of  $n$  identical cells would lead to  $n$  species, all differing only for the name of its inner compartment, instead of a single specie with multiplicity  $n$ . For circumventing this problem, in Section 4 we introduce a variant of the GSAM with a *copy-on-write* optimization strategy—hence called COWGSAM. The idea is to keep a single copy of each species, with its multiplicity; when a reaction has to be applied, fresh copies of the compartments involved are generated on-the-fly, and reactions and rates are updated accordingly. In this way, the hierarchical structure is unfolded only if and when needed.

In Section 5 we show how the Brane Calculus can be represented in the COWGSAM, and we will prove that the abstract machine obtained in this way is adequate with respect to the stochastic semantics of the Brane Calculus; in this proof, we take advantage of the compositional definition of this semantics.

Conclusions and final remarks are in Section 6.

## 2 Brane Calculus

In this section we recall Cardelli’s Brane Calculus [4] focusing on its basic version (without communication primitives, complexes and replication).

First, let us fix the notation we will use hereafter. Let  $S$  be a set of *sorts* (or “types”), ranged over by  $s, t$ , and  $T$  a set of  $S$ -sorted terms; for  $t \in S$ ,  $T_t \subseteq T$  denotes the set of terms of sort  $t$ . For  $A$  a set of symbols,  $A^*$  denotes the set of finite words (or lists) over  $A$ , and  $\langle a_1, \dots, a_n \rangle$  denotes a word in  $A^*$ . For a word  $\langle t_1, \dots, t_n \rangle$  in  $S^*$ , we define  $T_{\langle t_1, \dots, t_n \rangle} \triangleq T_{t_1} \times \dots \times T_{t_n}$ .

**Syntax** The sorts and the set  $\mathbb{B}$  of terms of Brane Calculus are the following:

$$\begin{array}{ll}
 \text{Sorts} :: S & t ::= \text{sys} \mid \text{mem} \\
 \text{Membranes} :: \mathbb{B}_{\text{mem}} & \sigma, \tau ::= \mathbf{0} \mid \sigma|\tau \mid \vartheta_n.\sigma \mid \vartheta_n^\perp(\tau).\sigma \mid \vartheta_n.\sigma \mid \vartheta_n^\perp.\sigma \mid \circlearrowleft_n(\tau).\sigma \\
 \text{Systems} :: \mathbb{B}_{\text{sys}} & P, Q ::= \diamond \mid P \circ Q \mid \sigma(P)
 \end{array}$$

The subscripted names  $n$  are taken from a countable set  $\Lambda$ . By convention we shall use  $M, N, \dots$  to denote generic Brane Calculus terms in  $\mathbb{B}$ .

A *membrane* can be either the empty membrane  $\mathbf{0}$ , or the parallel composition of two membranes  $\sigma|\tau$ , or the action-prefixed membrane  $\varepsilon.\sigma$ . Actions are: *phagocytosis*  $\vartheta$ , *exocytosis*  $\vartheta$ , and *pinocytosis*  $\circlearrowleft$ . Each action but pinocytosis comes with a matching co-action, indicated by the superscript  $^\perp$ .

A *system* can be either the empty system  $\diamond$ , or the parallel composition  $P \circ Q$ , or the system nested within a membrane  $\sigma(P)$ . Notice that, differently from [4], pino actions are indexed by names in  $\Lambda$ . In [4], names are meant only to pair up an action with its corresponding co-action, hence a pino action does not need to be indexed by any name. Actually, names can be thought of as an abstract representation

$$\begin{array}{c}
\frac{}{\vartheta_n^+(\rho).\tau|\tau_0(Q) \circ \vartheta_n.\sigma|\sigma_0(P) \twoheadrightarrow \tau|\tau_0(\rho(\sigma|\sigma_0(P))) \circ Q} \text{ (red-phago)} \\
\frac{}{\vartheta_n^+.\tau|\tau_0(\vartheta_n.\sigma|\sigma_0(P) \circ Q) \twoheadrightarrow \sigma|\sigma_0|\tau|\tau_0(Q) \circ P} \text{ (red-exo)} \\
\frac{}{\otimes(\rho).\sigma|\sigma_0(P) \twoheadrightarrow \sigma|\sigma_0(\rho(\diamond) \circ P)} \text{ (red-pino)} \quad \frac{P \twoheadrightarrow Q}{\sigma(P) \twoheadrightarrow \sigma(Q)} \text{ (red-loc)} \\
\frac{P \twoheadrightarrow Q}{P \circ R \twoheadrightarrow Q \circ R} \text{ (red-comp)} \quad \frac{P \equiv P' \quad P' \twoheadrightarrow Q' \quad Q' \equiv Q}{P \twoheadrightarrow Q} \text{ (red-equiv)}
\end{array}$$

Table 1: Reduction semantics for the Brane Calculus.

of particular protein conformational shapes; hence, each name can correspond to a different biological behaviour. Therefore, if we want to observe also kinetic properties of processes, it is important to keep track of names in pino actions.

Terms can be rearranged according to a structural congruence relation; the intended meaning is that two congruent terms actually denote the same system. Structural congruence  $\equiv$  is the smallest equivalence relation over  $\mathbb{B}$  which satisfies the axioms and rules listed below.

$$\begin{array}{c}
P \circ Q \equiv Q \circ P \quad P \circ (Q \circ R) \equiv (P \circ Q) \circ R \quad P \circ \diamond \equiv P \\
\sigma|\tau \equiv \tau|\sigma \quad \sigma|(\tau|\rho) \equiv (\sigma|\tau)|\rho \quad \sigma|\mathbf{0} \equiv \sigma \\
\mathbf{0}(\diamond) \equiv \diamond \quad \frac{P \equiv Q}{P \circ R \equiv Q \circ R} \quad \frac{\sigma \equiv \tau}{\sigma|\rho \equiv \tau|\rho} \quad \frac{P \equiv Q \quad \sigma \equiv \tau}{\sigma(P) \equiv \tau(Q)} \\
\frac{\alpha \in \{\vartheta_n, \vartheta_n^+, \vartheta_n^+\}_{n \in \Lambda} \quad \sigma \equiv \tau}{\alpha.\sigma \equiv \alpha.\tau} \quad \frac{\beta \in \{\vartheta_n^+, \otimes_n\}_{n \in \Lambda} \quad \rho \equiv \nu \quad \sigma \equiv \tau}{\beta(\rho).\sigma \equiv \beta(\nu).\tau}
\end{array}$$

Differently from [4], we allow to rearrange also the sub-membranes contained in co-phago and pino actions (by means of the last inference rule above).

**Reduction Semantics** The dynamic behaviour of Brane Calculus is specified by means of a reduction semantics, defined over a reduction relation (“reaction”)  $\twoheadrightarrow \subseteq \mathbb{B}_{\text{sys}} \times \mathbb{B}_{\text{sys}}$ , whose rules are listed in Table 1. Notice that the presence of (red-phago/exo/pino) and (red-equiv) makes this not a *structural* presentation, since these rules are not primitive recursive in the syntax (i.e., structural recursive) as required by the SOS format.

### 3 Stochastic Structural Operational Semantics for the Brane Calculus

In this section we recall the stochastic structural operational semantics for the Brane Calculus, as defined in [2]. Following [5], we replace the classic “pointwise” rules of the form  $P \xrightarrow{a,r} P'$  with rules of the form  $P \rightarrow \mu$ , where  $\mu$  is an indexed class of *measures* on the *measurable space* of processes. We assume the reader to be familiar with basic notions from measure theory; for a brief summary, see Appendix A.

$$\begin{array}{c}
\frac{}{\mathbf{0} \rightarrow_{\text{mem}} \omega_{\text{mem}}} \text{ (zero)} \quad \frac{\alpha \in \{\mathfrak{V}_n, \mathfrak{D}_n, \mathfrak{V}_n^\perp \mid n \in \Lambda\}}{\alpha \cdot \sigma \rightarrow_{\text{mem}} [\alpha]_\sigma} \text{ (pref)} \\
\frac{\beta \in \{\mathfrak{V}_n^\perp, \mathfrak{O}_n \mid n \in \Lambda\}}{\beta(\tau) \cdot \sigma \rightarrow_{\text{mem}} [\beta]_\sigma^\tau} \text{ (pref-arg)} \quad \frac{\sigma \rightarrow_{\text{mem}} \mu' \quad \tau \rightarrow_{\text{mem}} \mu''}{\sigma | \tau \rightarrow_{\text{mem}} \mu' \sigma \oplus_\tau \mu''} \text{ (par)} \\
\frac{}{\diamond \rightarrow_{\text{sys}} \omega_{\text{sys}}} \text{ (void)} \quad \frac{\sigma \rightarrow_{\text{mem}} \nu \quad P \rightarrow_{\text{sys}} \mu}{\sigma(P) \rightarrow_{\text{sys}} \mu \otimes_P^\sigma \nu} \text{ (loc)} \quad \frac{P \rightarrow_{\text{sys}} \mu' \quad Q \rightarrow_{\text{sys}} \mu''}{P \circ Q \rightarrow_{\text{sys}} \mu' P \otimes_Q \mu''} \text{ (comp)}
\end{array}$$

Table 2: Stochastic structural operational semantics for Brane Calculus

The set of action labels for the Brane Calculus will be denoted by  $\mathbb{A}$  and can be partitioned with respect to the source sort (i.e., either systems or membranes), as follows:

$$\begin{aligned}
\mathbb{A}_{\text{sys}} &\triangleq \{id : \text{sys} \rightarrow \langle \text{sys} \rangle\} \cup \{\text{ph}_n : \text{sys} \rightarrow \langle \text{sys}, \text{sys} \rangle \mid n \in \Lambda\} \cup \\
&\quad \{\text{ph}_n^\perp : \text{sys} \rightarrow \langle \text{mem}, \text{mem}, \text{sys}, \text{sys} \rangle \mid n \in \Lambda\} \cup \\
&\quad \{\text{ex}_n : \text{sys} \rightarrow \langle \text{mem}, \text{sys}, \text{sys} \rangle \mid n \in \Lambda\} \\
\mathbb{A}_{\text{mem}} &\triangleq \{\mathfrak{V}_n, \mathfrak{D}_n, \mathfrak{V}_n^\perp : \text{mem} \rightarrow \langle \text{mem} \rangle \mid n \in \Lambda\} \cup \\
&\quad \{\mathfrak{V}_n^\perp, \mathfrak{O}_n : \text{mem} \rightarrow \langle \text{mem}, \text{mem} \rangle \mid n \in \Lambda\}
\end{aligned}$$

Let  $a$  range over  $\mathbb{A}$ , and  $ar(a)$  denote its arity. To ease the reading in the following we will use the notation  $\Delta_a(T, \Sigma)$  to denote the set of measures  $\Delta(T_{\langle t_1, \dots, t_n \rangle}, \bigotimes_{i=1}^n \Sigma_{t_i})$ , for  $ar(a) = t \rightarrow \langle t_1, \dots, t_n \rangle$ .

Let  $\mathbb{B}/\equiv$  be the set of  $\equiv$ -equivalence classes on  $\mathbb{B}$ . For  $M \in \mathbb{B}$ , we denote by  $[M]_{\equiv}$  the  $\equiv$ -equivalence class of  $M$  (sometimes dropping the equivalence symbol when clear from the context).

**Definition 3.1 (Measurable space of terms).** *The measurable space of terms  $(\mathbb{B}, \Pi)$  is given by the measurable space over  $\mathbb{B}$  where  $\Pi$  is the  $\sigma$ -algebra generated by  $\mathbb{B}/\equiv$ .*

Notice that  $\mathbb{B}/\equiv$  is a denumerable partition of  $\mathbb{B}$ , hence it is a base (a generator such that all its elements are disjoint) for  $\Pi$ . Any element of  $\Pi$  can be obtained by a countable union of elements of the base, i.e., for all  $\mathcal{M} \in \Pi$  there exist  $\{M_i\}_{i \in I}$ , for some countable  $I$ , such that  $\mathcal{M} = \bigcup_{i \in I} [M_i]_{\equiv}$ . As a consequence, in order to generate the whole  $\Pi$  we can simply compute all these unions, without the need of any closure by complement.

A similar argument holds for the product space  $(\mathbb{B}_{\langle t_1, \dots, t_n \rangle}, \bigotimes_{i=1}^n \Pi_{t_i})$ , where  $t_i \in \{\text{mem}, \text{sys}\}$  ( $1 \leq i \leq n$ ); indeed  $\bigotimes_{i=1}^n \Pi_{t_i}$  can be generated from the base  $\mathbb{B}_{\langle t_1, \dots, t_n \rangle} / \equiv_{\langle t_1, \dots, t_n \rangle}$ , where  $\equiv_{\langle t_1, \dots, t_n \rangle} \subseteq \mathbb{B}_{\langle t_1, \dots, t_n \rangle} \times \mathbb{B}_{\langle t_1, \dots, t_n \rangle}$  is defined by

$$\langle M_1, \dots, M_n \rangle \equiv_{\langle t_1, \dots, t_n \rangle} \langle N_1, \dots, N_n \rangle \quad \text{iff} \quad M_i \equiv N_i, \text{ for all } 1 \leq i \leq n,$$

which can be easily checked to be an equivalence relation.  $\equiv_{\langle t_1, \dots, t_n \rangle}$ -equivalence classes are rectangles, i.e.  $[\langle M_1, \dots, M_n \rangle]_{\equiv_{\langle t_1, \dots, t_n \rangle}} = [M_1]_{\equiv} \times \dots \times [M_n]_{\equiv}$ , therefore the product measure  $\bigotimes_{i=1}^n \Pi_{t_i}$  is well defined. For sake of simplicity in the following we write  $[\langle M_1, \dots, M_n \rangle]_{\equiv}$  in place of  $[\langle M_1, \dots, M_n \rangle]_{\equiv_{\langle t_1, \dots, t_n \rangle}}$ , and  $\mathbb{B}_{\langle t_1, \dots, t_n \rangle} / \equiv$  in place of  $\mathbb{B}_{\langle t_1, \dots, t_n \rangle} / \equiv_{\langle t_1, \dots, t_n \rangle}$ .

The operational semantics associates with each membrane a family of measures in  $\Delta^{\mathbb{A}_{\text{mem}}}(\mathbb{B}, \Pi)$ , and with each system a family of measures in  $\Delta^{\mathbb{A}_{\text{sys}}}(\mathbb{B}, \Pi)$ . This can be represented by two relations

$\rightarrow_{\text{mem}}: T_{\text{mem}} \rightarrow \Delta^{\mathbb{A}_{\text{mem}}}(\mathbb{B}, \Pi)$ ,  $\rightarrow_{\text{sys}}: T_{\text{sys}} \rightarrow \Delta^{\mathbb{A}_{\text{sys}}}(\mathbb{B}, \Pi)$ , defined by the SOS rules listed in Table 2. (In the following, for sake of readability, we will drop the indexes  $_{\text{mem}, \text{sys}}$ ). In these rules we use some constants and operations over indexed families of measures, that we define next. For a set (of labels)  $A$ , let us denote by  $\Delta^A(\mathbb{B}, \Pi)$  the set  $\prod_{a \in A} \Delta_a(\mathbb{B}, \Pi)$  of  $A$ -indexed families of measures over  $(\mathbb{B}, \Pi)$ . Given a family of measures  $\mu \in \Delta^A(\mathbb{B}, \Pi)$  and  $a \in A$ , the  $a$ -component of  $\mu$  will be denoted as  $\mu_a \in \Delta_a(\mathbb{B}, \Pi)$ .

**Null:** Let  $\omega_{\text{mem}} \in \Delta^{\mathbb{A}_{\text{mem}}}(\mathbb{B}, \Pi)$  be the constantly zero measure, i.e., for all  $a \in \mathbb{A}_{\text{mem}}$  such that  $ar(a) = t \rightarrow \langle t_1, \dots, t_n \rangle$  and  $\mathcal{M} \in \otimes_{i=1}^n \Pi_{t_i}$ :  $(\omega_{\text{mem}})_a(\mathcal{M}) = 0$ .

**Prefix:** For arbitrary  $n \in \Lambda$ ,  $\alpha \in \{\wp, \wp, \wp^\perp\}$ , and  $\beta \in \{\wp^\perp, \odot\}$ , let the constants  $[\alpha_n]_\sigma, [\beta_n]_\sigma^\tau \in \Delta^{\mathbb{A}_{\text{mem}}}(\mathbb{B}, \Pi)$  be defined, for arbitrary  $X, Y \in \mathbb{B}_{\text{mem}}/\equiv$ , as follows:

$$\begin{aligned} ([\alpha_n]_\sigma)_{\alpha_m}(X) &= \begin{cases} \iota(n) & \text{if } n = m \text{ and } \sigma \in X \\ 0 & \text{otherwise} \end{cases} \\ ([\alpha_n]_\sigma)_{\beta_m}(X \times Y) &= 0 \\ ([\beta_n]_\sigma)_{\alpha_m}(X) &= 0 \\ ([\beta_n]_\sigma^\tau)_{\beta_m}(X \times Y) &= \begin{cases} \iota(n) & \text{if } n = m \text{ and } \sigma \in X, \tau \in Y \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

**Parallel:** For  $\mu, \mu' \in \Delta^{\mathbb{A}_{\text{mem}}}(\mathbb{B}, \Pi)$ , let  $\mu_\sigma \odot_\tau \mu' \in \Delta^{\mathbb{A}_{\text{mem}}}(\mathbb{B}, \Pi)$  be defined, for  $n \in \Lambda$ ,  $\alpha \in \{\wp, \wp, \wp^\perp\}$ ,  $\beta \in \{\wp^\perp, \odot\}$ , and  $X, Y \in \mathbb{B}_{\text{mem}}/\equiv$ , as follows (where for all  $X, \tau$ :  $X|_\tau \triangleq \bigcup \{[\sigma]_{\equiv} \mid \sigma|_\tau \in X\}$ ):

$$\begin{aligned} (\mu_\sigma \odot_\tau \mu')_{\alpha_n}(X) &= \mu_{\alpha_n}(X|_\tau) + \mu'_{\alpha_n}(X|_\sigma) \\ (\mu_\sigma \odot_\tau \mu')_{\beta_n}(X \times Y) &= (\mu)_{\beta_n}(X|_\tau \times Y) + (\mu')_{\beta_n}(X|_\sigma \times Y) \end{aligned}$$

**Void:** Let  $\omega_{\text{sys}} \in \Delta^{\mathbb{A}_{\text{sys}}}(\mathbb{B}, \Pi)$  be defined by  $(\omega_{\text{sys}})_a(\mathcal{M}) = 0$  for any  $a \in \mathbb{A}_{\text{sys}}$ , such that  $ar(a) = t \rightarrow \langle t_1, \dots, t_n \rangle$ , and  $\mathcal{M} \in \otimes_{i=1}^n \Pi_{t_i}$ .

**Nesting:** For  $\nu \in \Delta^{\mathbb{A}_{\text{mem}}}(\mathbb{B}, \Pi)$  and  $\mu \in \Delta^{\mathbb{A}_{\text{sys}}}(\mathbb{B}, \Pi)$ , let  $\mu @_{\mathcal{P}}^\sigma \nu \in \Delta^{\mathbb{A}_{\text{sys}}}(\mathbb{B}, \Pi)$  be defined, for  $X, Y \in \mathbb{B}_{\text{mem}}/\equiv$  and  $Z, W \in \mathbb{B}_{\text{sys}}/\equiv$ , as follows:

$$\begin{aligned} (\mu @_{\mathcal{P}}^\sigma \nu)_{\text{ph}_n}(Z \times W) &= \begin{cases} \nu_{\wp_n}([\sigma]_{\equiv}) & \text{if } \sigma(\mathcal{P}) \in Z \text{ and } \diamond \in W \\ 0 & \text{otherwise} \end{cases} \\ (\mu @_{\mathcal{P}}^\sigma \nu)_{\text{ph}_n^\perp}(X \times Y \times Z \times W) &= \begin{cases} \nu_{\wp_n^\perp}(X \times Y) & \text{if } \mathcal{P} \in Z \text{ and } \diamond \in W \\ 0 & \text{otherwise} \end{cases} \\ (\mu @_{\mathcal{P}}^\sigma \nu)_{\text{ex}_n}(X \times Z \times W) &= \begin{cases} \nu_{\wp_n}(X) & \text{if } \mathcal{P} \in Z \text{ and } \diamond \in W \\ 0 & \text{otherwise} \end{cases} \\ (\mu @_{\mathcal{P}}^\sigma \nu)_{\text{id}}(X) &= \mu_{\text{id}}(X_{\sigma(\mathcal{D})}) + \sum_{\substack{n \in \Lambda \\ X' \mathcal{D} X'' \mathcal{D} [\diamond]_{\equiv} \mathcal{D} \circ [P]_{\equiv} \mathcal{D} = X}}^{n \in \Lambda} \nu_{\wp_n}(X' \times X'') + \\ &\quad \sum_{\substack{n \in \Lambda \\ X' | X'' \mathcal{D} (Y'') \mathcal{D} \circ Y' = X}} \frac{\mu_{\text{ex}_n}(X' \times Y' \times Y'') \cdot \nu_{\wp_n^\perp}(X'')}{\iota(n)} \end{aligned}$$

**Composition:** For  $\mu, \mu' \in \Delta^{\text{A}_{\text{sys}}}(\mathbb{B}, \Pi)$ , let  $\mu_P \otimes_Q \mu' \in \Delta^{\text{A}_{\text{sys}}}(\mathbb{B}, \Pi)$  be defined, for  $X, Y \in \mathbb{B}_{\text{mem}/\equiv}$  and  $Z, W \in \mathbb{B}_{\text{sys}/\equiv}$ , as follows (where for all  $W, Q, W \circ_Q \triangleq \bigcup \{ [P]_{\equiv} \mid P \circ Q \in W \}$ ):

$$\begin{aligned}
(\mu_P \otimes_Q \mu')_{\text{ph}_n}(Z \times W) &= \mu_{\text{ph}_n}(Z \times W \circ_Q) + \mu'_{\text{ph}_n}(Z \times W \circ_P) \\
(\mu_P \otimes_Q \mu')_{\text{ph}_n^\perp}(X \times Y \times Z \times W) &= \mu_{\text{ph}_n^\perp}(X \times Y \times Z \times W \circ_Q) + \mu'_{\text{ph}_n^\perp}(X \times Y \times Z \times W \circ_P) \\
(\mu_P \otimes_Q \mu')_{\text{ex}_n}(X \times Z \times W) &= \mu_{\text{ex}_n}(X \times Z \times W \circ_Q) + \mu'_{\text{ex}_n}(X \times Z \times W \circ_P) \\
(\mu_P \otimes_Q \mu')_{\text{id}}(X) &= \mu_{\text{id}}(X \circ_Q) + \mu'_{\text{id}}(X \circ_P) + \\
&\quad \sum_{\substack{n \in \Lambda \\ X_1 \Downarrow X_2 \Downarrow Y_1 \Downarrow Z_1 \Downarrow Y_2 \circ Z_2 = X}} \frac{\mu_{\text{ph}_n}(Y_1 \times Y_2) \cdot \mu'_{\text{ph}_n^\perp}(X_1 \times X_2 \times Z_1 \times Z_2)}{t(n)} + \\
&\quad \sum_{\substack{n \in \Lambda \\ X_1 \Downarrow X_2 \Downarrow Z_1 \Downarrow Y_1 \Downarrow Z_2 \circ Y_2 = X}} \frac{\mu_{\text{ph}_n^\perp}(X_1 \times X_2 \times Y_1 \times Y_2) \cdot \mu'_{\text{ph}_n}(Z_1 \times Z_2)}{t(n)}
\end{aligned}$$

These operators have nice algebraic properties (e.g.,  $\mu' \sigma \oplus \tau \mu'' = \mu'' \tau \oplus \sigma \mu'$ ,  $(\mu' \sigma \oplus \tau \mu'')_{\sigma | \tau \oplus \rho} \mu''' = \mu' \sigma \oplus \tau | \rho (\mu'' \tau \oplus \rho \mu''')$ , ...), and respect the structural congruence (e.g., if  $P \equiv P'$  and  $Q \equiv Q'$  then  $\mu'_{P \otimes Q} \mu'' = \mu'_{P' \otimes Q'} \mu''$ ). We refer to [2] for further details about these properties, very useful in calculations.

The next lemmata state that the stochastic transition relation  $\rightarrow$  (and hence the operational semantics) is well-defined and consistent, that is, for each process we have exactly one family of measures of its continuations, and this family respects structural congruence.

**Lemma 3.2 (Uniqueness).** *For  $a \in \mathbb{A}$  such that  $ar(a) = t \rightarrow \langle t_1, \dots, t_n \rangle$ , and  $M \in \mathbb{B}_t$ , there exists a unique  $\mu \in \Delta^{\mathbb{A}_t}(\mathbb{B}, \Pi)$  such that  $M \rightarrow \mu$ .*

**Lemma 3.3.** *If  $M \equiv N$  and  $M \rightarrow \mu$ , then  $N \rightarrow \mu$ .*

This operational semantics can be used to define the “traditional” pointwise semantics:

$$M \xrightarrow{a,r} \langle M_1, \dots, M_n \rangle \xleftrightarrow{\Delta} M \rightarrow \mu \text{ and } \mu_a([\langle M_1, \dots, M_n \rangle]_{\equiv}) = r$$

and it is conservative with respect to the non-stochastic reduction semantics.

**Proposition 3.4.** *For all systems  $P, Q \in \mathbb{B}_{\text{mem}}$ , if  $P \rightarrow \mu$  and  $\mu_{\text{id}}([Q]) > 0$  then  $P \twoheadrightarrow Q$ .*

## 4 The COW Generic Stochastic Abstract Machine

In this section we present a variant of the generic stochastic abstract machine (GSAM), oriented to systems with nested compartments.

The GSAM has been introduced in [13, 10] for simulating a broad range of process calculi with an arbitrary reaction-based simulation algorithm. Although it does not have a native notion of “compartment”, nested systems can be represented by “flattening” all compartments and species into a single multiset, where each species is tagged with names representing their position in the hierarchy, as shown in [10]. The idea is to represent each compartment as a different species, keeping track of their position in the hierarchy by means of (*node*) *names*. These names are ranged over by  $x, y, z, \dots$ , and are different from names in actions. As an example, a system  $\sigma(\tau(\mathbb{D}))$  is represented as the multiset  $\{\sigma(\mathbb{D}_y^x) \mapsto 1, \tau(\mathbb{D}_z^y) \mapsto 1\}$ , which means “there is one cell with membrane  $\sigma$  located in the compartment  $x$  and whose compartment is  $y$ , and one cell with membrane  $\tau$  positioned in  $y$  and whose compartment is  $z$ ”. Reactions can happen only if the names tagging the involved species match according to the required nesting structure.

Unfortunately, this approach does not scale well, as the population of species grows. Let us consider a system composed of  $n$  copies of the same cell, e.g.,  $n \cdot (\sigma\mathbb{D})$  (where  $n$  can be easily in the order of  $10^3$ – $10^4$ ). In the original GSAM idea, this should be represented in the machine as a single species with multiplicity  $n$ , and each possible reaction is represented once but with propensity given by the law of mass action taking into account the species' multiplicity  $n$ . Instead, the “flat encoding” above yields  $n$  different species  $\sigma\mathbb{C}_{y_1}^x\mathbb{D}, \dots, \sigma\mathbb{C}_{y_n}^x\mathbb{D}$ , each with multiplicity 1; the set of reactions explodes correspondingly.

For circumventing this problem we introduce a variant of the GSAM with a *copy-on-write* strategy—hence it is called COWGSAM. The idea is to keep a single copy of each species, with its multiplicity; the same applies to reactions. When a reaction has to be applied, the compartments involved are “unfolded”, i.e., fresh copies of the compartments are generated and the reaction set is modified accordingly; then, the reaction can be applied. In this way, the hierarchical structure is unfolded only if and when needed.

In order to implement this idea, we have to modify the notion of machine term, reaction and reaction rule. The COWGSAM (with the Next Reaction method) is shown in Figure 1.

First, for generating fresh names, we have to keep track of those already allocated. To this end we introduce *environments*, which are finite sets of names. Then, the machine state is represented by a *machine term*  $T$ , i.e. a quadruple  $E \vdash (t, S, R)$  where  $E$  is an environment;  $t$  is the current time;  $S$  is a finite function mapping each species  $I$  to its population  $S(I)$  (if not null); and  $R$  maps each reaction  $O$  to its activity  $A$ , which is used to compute the next reaction. (Notice that the syntax of species  $I$  is left unspecified, as it depends on the specific process calculus one has to implement.) We say that a machine term  $E \vdash (t, S, R)$  is *well-formed* if for all  $x_i, x_j \in E : x_i = x_j \Rightarrow i = j$ , and the free names in  $S, R$  appear in  $E$ . In the following, we assume that machine terms are well formed.

Each reaction is a quadruple  $(S_1, r, f, S_2)$ , basically representing a reaction  $S_1 \xrightarrow{r} S_2$ , where

- $S_1$  and  $S_2$  denote the *reactant* population and *product* population respectively;
- $r$  denotes the rate (in  $s^{-1}$ ) of the reaction;
- $f$  is a function mapping machine terms to machine terms; this functions implements any global update of the machine term after the reaction (if needed).

The *transitions* of the abstract machine are represented by a relation  $T \xrightarrow{a, O} T'$  between machine terms, indexed by the propensity  $a$  and reaction rules. This should be read as “ $T$  goes to  $T'$  with rate  $a$ , using the rule  $O$ ”. This relation is defined by (Reaction rule) in Figure 1, where the function  $next(T)$  selects the next reaction, i.e. it returns a pair  $(O, t')$  where  $O = (S_1, r, f, S_2)$  is the reaction to happen first among all possible reactions in  $T$ , and  $t'$  is the new time of the system. Once the reaction has been selected, we have first to create the separate (private) copies of the compartments involved, and to update the reaction set accordingly. This is done by the functions  $cow(-)$  and  $dup(-)$ , which implement a deep copy-on-write:  $cow(E \vdash (t, S, R), S_1)$  is a machine term  $E' \vdash (t, S', R')$  representing the same state as  $E \vdash (t, S, R)$ , but in  $S'$  the species indicated in  $S_1$  are unfolded;  $E'$  contains all names which have been generated in the process, and  $R'$  is the new set of reactions. (Actually, the freshly generated copies represent the instances which are *not* involved by the reaction; this simplifies the reaction application.) An example of the action of  $cow(-)$  is depicted in Figure 2.

At this point, the reaction is executed, by removing the reactants  $S_1$  from the machine term (via the operation  $\ominus$ ), adding the products  $S_2$  (via the operation  $\oplus$ ) and updating the current time of the machine. The function  $f$  performs any global “clean-up” and restructuring to the machine term that may be required by the reaction (e.g., garbage collection, elimination of names not used anymore, . . .). Moreover, since a reaction can rearrange the hierarchy structure, possibly creating new compartments and deleting others, we have to add to the environment any fresh name introduced in the products.

$$\begin{array}{ll}
T ::= E \vdash (t, S, R) & \text{(Machine term)} \\
E ::= x_1, \dots, x_n & \text{(Environments)} \\
S ::= \{I_1 \mapsto i_1, \dots, I_N \mapsto i_N\} & \text{(Populations)} \\
R ::= \{O_1 \mapsto A_1, \dots, O_N \mapsto A_N\} & \text{(Reactions)} \\
O ::= (S_1, r, f, S_2) & \text{(Reaction)}
\end{array}$$

$$\frac{((S_1, r, f, S_2), a, t') = \text{next}(t, S, R) \quad E' \vdash (t, S', R') = \text{cow}(E \vdash (t, S, R), S_1)}{E \vdash (t, S, R) \xrightarrow{a, (S_1, r, f, S_2)} \text{norm}(E' \cup \text{fn}(S_2) \vdash (f(S_2 \oplus ((t', S', R') \ominus S_1))))} \quad \text{(Reaction rule)}$$

$$\begin{aligned}
\text{next}(t, S, R) &\triangleq (O, a, t') \quad \text{if } R(O) = (t', a) \text{ and } t' = \min\{t \mid R(O) = (t, a)\} \\
\text{cow}(E \vdash (t, S, R), \emptyset) &\triangleq E \vdash (t, S, R) \\
\text{cow}(E \vdash (t, S, R), \{\rho \mathcal{Q}_z^y \mathcal{D} \mapsto j\} \cup S_1) &\triangleq \text{cow}(E' \cup \text{fn}(S'') \vdash (t, S'', R'' \cup \text{init}(L' \cup L'', (t, S', R))), S_1) \quad \text{where} \\
&\text{if } S(\sigma \mathcal{Q}_y^x \mathcal{D}) > 1 \text{ for some } \sigma, x : \\
&\quad \text{let } i = S(\sigma \mathcal{Q}_y^x \mathcal{D}) \text{ and } y' \notin E \\
&\quad (S', R') = \text{dup}(E \vdash (t, S\{\sigma \mathcal{Q}_y^x \mathcal{D} \mapsto 1, \sigma \mathcal{Q}_y^x \mathcal{D} \mapsto i-1\}, R), y, y') \\
&\quad L' = \text{reactions}(\sigma \mathcal{Q}_y^x \mathcal{D} \mapsto i-1, S) \\
&\quad \text{otherwise let } (S', R') = (S, R), L' = \emptyset \text{ in} \\
&\quad \text{let } n = S(\rho \mathcal{Q}_z^y \mathcal{D}), E' = E \cup \text{fn}(S'), \text{ and } z' \notin E' \\
&\quad (S'', R'') = \text{dup}(E' \vdash (t, S'\{\rho \mathcal{Q}_z^y \mathcal{D} \mapsto j, \rho \mathcal{Q}_z^y \mathcal{D} \mapsto n-j\}, R'), z, z') \\
&\quad L'' = \text{reactions}(\rho \mathcal{Q}_z^y \mathcal{D} \mapsto n-j, S') \\
\text{dup}(E \vdash (t, S, R), y, y') &\triangleq (S' \cup S'', R \cup R'' \cup \text{init}(L, (t, S', R))) \\
&\text{where } S' = S \cup \{\rho \mathcal{Q}_w^y \mathcal{D} \mapsto i \mid S(\sigma \mathcal{Q}_y^z \mathcal{D}) > 0, S(\rho \mathcal{Q}_w^y \mathcal{D}) = i\}, w' \notin E, \\
&\quad L = \text{reactions}(\rho \mathcal{Q}_w^y \mathcal{D} \mapsto i, S'), \\
&\quad (S'', R'') = \text{dup}(E \cup \{w'\} \vdash (t, S, R), w, w')
\end{aligned}$$

Figure 1: The COW Generic Stochastic Abstract Machine, with the Next Reaction method.

Finally, the term can be “normalized” by collapsing equivalent copies of the same subtree into a single copy (with the right multiplicity), by the function  $\text{norm}(\cdot)$ . In its simplest form,  $\text{norm}(\cdot)$  can be the identity, i.e., no normalization is performed at all. Although this is correct, it can lead to unnecessary copies of the same subtrees. We can define  $\text{norm}(\cdot)$  such that

$$\text{norm}(\{I_1 \mapsto i_1, I_2 \mapsto i_2\} \cup S, R) = \text{norm}(\{I_1 \mapsto i_1 + i_2\} \cup S, R[I_2/I_1]) \quad \text{if } I_1 \equiv I_2$$

where the equivalence between species can be implemented by comparing the subtrees starting from  $I_1$ ,  $I_2$ , e.g., by calculating a suitable hash value. We leave this refinement as future development.

In order to implement the Next Reaction method, each reaction  $O$  is associated with a pair  $R(O) = (a, t)$ , where  $a$  is the reaction *propensity* and  $t$  is the time at which the reaction is supposed to occur. The function  $\text{delay}(r, a)$  computes a time interval from a random variable with rate  $r$  and propensity  $a$ .

$$\begin{aligned}
\{I_1 \mapsto i_1, \dots, I_N \mapsto i_N\} \oplus (t, S, R) &\triangleq I_1 \mapsto i_1 \oplus \dots \oplus I_N \mapsto i_N \oplus (t, S, R) \\
I \mapsto i \oplus (t, S, R) &\triangleq \begin{cases} (t, S', R \cup \text{updates}(I, (t, S', R))) & \text{if } S(I) = i' \text{ and } S' = S\{I \mapsto i' + i\} \\ (t, S', R \cup \text{init}(L, (t, S', R))) & \text{if } I \notin \text{dom}(S), S' = S\{I \mapsto i\} \\ & \text{and } L = \text{reactions}(I \mapsto i, S) \end{cases} \\
(t, S, R) \ominus \{I_1 \mapsto i_1, \dots, I_N \mapsto i_N\} &\triangleq (t, S, R) \ominus I_1 \mapsto i_1 \ominus \dots \ominus I_N \mapsto i_N \\
(t, S, R) \ominus I \mapsto i &\triangleq (t, S', R \cup \text{updates}(I, (t, S', R))) \\ &\quad \text{if } S(I) = i', i' \geq i \text{ and } S' = S\{I \mapsto i' - i\} \\
\text{init}(L, (t, S, R)) &\triangleq \{O \mapsto (t', a) \mid O \in L \text{ and } a = \text{propensity}(O, S) \text{ and} \\ &\quad O = (S_1, r, f, S_2) \text{ and } t' = t + \text{delay}(r, a)\} \\
\text{updates}(I, (t, S, R)) &\triangleq \{O \mapsto (t', a') \mid R(O) = (t'', a) \text{ and } O = (S_1, r, f, S_2) \text{ and } S_1(I) > 0 \text{ and} \\ &\quad \text{if } t'' > t \text{ then } a' = \text{propensity}(O, S) \text{ and } t' = t + (a/a')(t'' - t) \\ &\quad \text{if } t'' = t \text{ then } a' = \text{propensity}(O, S) \text{ and } t' = t + \text{delay}(r, a)\} \\
\text{propensity}((S_1, r, f, S_2), S) &\triangleq r \cdot \binom{S^*(I_1)}{j_1} \cdot \dots \cdot \binom{S^*(I_n)}{j_n} \text{ if } S_1 = \{I_1 \mapsto j_1, \dots, I_n \mapsto j_n\} \\
S^*(\sigma \mathbb{Q}_y^x) &\triangleq \begin{cases} S(\sigma \mathbb{Q}_y^x) & \text{if } x = \text{root} \\ S(\sigma \mathbb{Q}_y^x) \cdot S^*(\rho \mathbb{Q}_x^z) & \text{if } x \neq \text{root} \text{ and } S(\rho \mathbb{Q}_x^z) > 0 \end{cases}
\end{aligned}$$

Figure 1: The COW Generic Stochastic Abstract Machine (cont.).

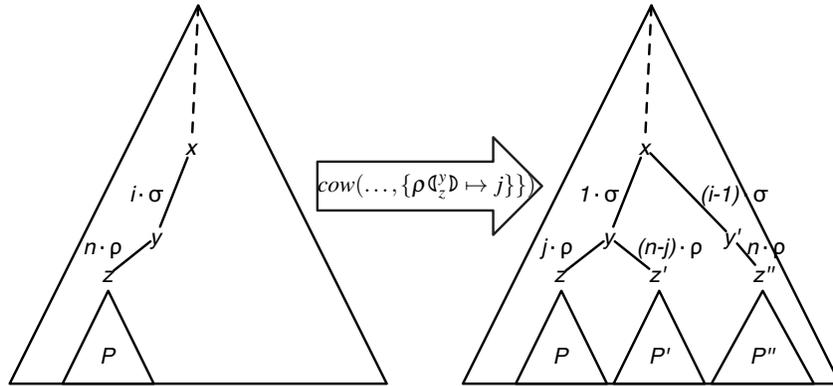


Figure 2: Example copy-on-write of a subterm.

This general structure can be instantiated with a given process calculus, just by providing the definition for the missing parts. Given a set  $Proc$  of processes, we have to define:

1. the syntax of the species  $I$  (which may be different from that of processes);
2. a function  $species_{E,x}(P)$  mapping a process  $P \in Proc$  to a species set located in  $x$ ;
3. a function  $reactions(I \mapsto i, S)$  for computing the multiset of reactions between a (new) species  $I$  with multiplicity  $i$  and a population of (existing) species  $S$ .

The function *species* is used to initialise the abstract machine at the beginning of a simulation. If we aim to simulate the execution of a process  $P \in Proc$ , the corresponding initial state (rooted in *root*) is

$$\llbracket P \rrbracket_{root} \triangleq fn(J) \vdash J \oplus (0, \emptyset, \emptyset) \quad \text{where } J = species_{\emptyset, root}(P).$$

The *reactions* function is used to adjust the set of possible reactions dynamically.

## 5 Implementing the Stochastic Brane Calculus in COWGSAM

In this section, we show how the COW Generic Stochastic Abstract Machine can be used to implement the Stochastic Brane Calculus, following the protocol described in the previous section.

### 5.1 Encoding of the Stochastic Brane Calculus

**Syntax of species** We define the species for the brane calculus, which in turn lead us to introduce *complexes* and *actions*. Notice that (despite the deceiving syntax) species are not systems and complexes are not membranes; nevertheless, actions are a subset of membranes.

$$\begin{aligned} I &::= C \langle \mathcal{D}_y^x \rangle && \text{(Species)} \\ C &::= A_1, \dots, A_n && \text{(Complexes)} \\ A &::= \vartheta_n \cdot \sigma \mid \vartheta_n^\perp(\tau) \cdot \sigma \mid \vartheta_n \cdot \sigma \mid \vartheta_n^\perp \cdot \sigma \mid \otimes_n(\tau) \cdot \sigma && \text{(Actions)} \end{aligned}$$

Node names can appear in the species in  $S$  and in reactions  $R$ .

**The species function** We can now provide the definition of the translation of a process  $P \in \mathbb{B}_{sys}$  into a species set. Basically, each compartment is assigned a different, fresh node name; therefore, the function  $species_{E,x}(P)$  is parametric in the set  $E$  of allocated node names, and the name  $x \in E$  to be used as the location of the system  $P$ . The name  $x$  changes as we descend the compartment hierarchy.

In order to capture correctly the multiplicity of each species, we assume that systems are in *normal form*. Basically, this form is a shorthand for products where  $n$  copies of the same system, i.e.,  $P \circ \dots \circ P$ , are represented as  $n \cdot P$ .

$$\begin{aligned} \text{Normal Systems } &:: \mathbb{B}_{sys}^n & Q &::= n_1 \cdot \sigma_1 \langle \mathcal{Q}_1 \rangle \circ \dots \circ n_k \cdot \sigma_k \langle \mathcal{Q}_k \rangle \\ &&& \text{where } k \geq 0 \text{ and for } i \neq j : \sigma_i \langle \mathcal{Q}_i \rangle \neq \sigma_j \langle \mathcal{Q}_j \rangle \end{aligned}$$

A system in normal form can be translated into a system just by unfolding the products. For  $Q \in \mathbb{B}_{sys}^n$  a system in normal form, let  $\lceil Q \rceil \in \mathbb{B}_{sys}$  defined as follows:

$$\lceil \diamond \rceil = \diamond \quad \lceil n \cdot \sigma \langle \mathcal{Q}' \rangle \circ Q'' \rceil = \overbrace{\sigma \langle \lceil \mathcal{Q}' \rceil \rangle \circ \dots \circ \sigma \langle \lceil \mathcal{Q}' \rceil \rangle}^{n \text{ times}} \circ \lceil Q'' \rceil$$

**Proposition 5.1.** *For all  $P \in \mathbb{B}_{sys}$ , there exists a system in normal form  $Q \in \mathbb{B}_{sys}^n$  such that  $\lceil Q \rceil \equiv P$ .*

As a consequence, we can give the definition of *species* on systems in normal form, as follows:

$$\begin{aligned} species_{E,x}(\diamond) &\triangleq \emptyset \\ species_{E,x}(n \cdot \sigma \langle \mathcal{Q}_1 \rangle \circ \mathcal{Q}_2) &\triangleq \{s(\sigma) \langle \mathcal{D}_y^x \rangle \mapsto n\} \cup species_{E \uplus \{y\}, y}(\mathcal{Q}_1) \cup species_{E,x}(\mathcal{Q}_2) \\ &\text{with } y \notin E \text{ and } fn(species_{E \uplus \{y\}, y}(\mathcal{Q}_1)) \cap fn(species_{E,x}(\mathcal{Q}_2)) \subseteq \{x\} \end{aligned}$$

The condition in the second case ensures that two different compartments are never given the same name—any name clash has to be resolved by an  $\alpha$ -conversion. The function  $s(\_)$  converts a membrane into a set of complexes:

$$s(\mathbf{0}) \triangleq \emptyset \quad s(\sigma|\sigma') \triangleq s(\sigma) \cup s(\sigma') \quad s(\pi.\sigma) \triangleq \{\pi.\sigma\}$$

**The reactions function** The next step is to define the function  $reactions(I \mapsto i, S)$ , for  $I$  a species with multiplicity  $i$  and  $S$  a population.

$$\begin{aligned} reactions_E(I_1 \mapsto i, S) &\triangleq unary_E(I_1) \cup binary_E(I_1, S) \\ unary_E(I_1) &\triangleq \{(\{I_1 \mapsto 1\}, r_n, id, \{(s(\sigma) \cup U'_1) \mathbb{Q}_y^x \mapsto 1, s(\rho) \mathbb{Q}_w^y \mapsto 1\}) \mid \\ &\quad I_1 = U_1 \mathbb{Q}_y^x, U_1 = \{\otimes_n(\rho).\sigma\} \cup U'_1, w \notin E\} \\ binary_E(I_1, S) &\triangleq \{(\{I_1 \mapsto 1, I_2 \mapsto 1\}, r_n, f, \{(s(\tau) \cup U'_1 \cup s(\sigma) \cup U'_2) \mathbb{Q}_y^x \mapsto 1\}) \mid I_2 \in dom(S), \\ &\quad (I_1 = U_1 \mathbb{Q}_y^x, I_2 = U_2 \mathbb{Q}_w^y) \vee (I_2 = U_1 \mathbb{Q}_y^x, I_1 = U_2 \mathbb{Q}_w^y), \\ &\quad U_1 = \{\otimes_n^{\perp}.\tau\} \cup U'_1, U_2 = \{\otimes_n.\sigma\} \cup U'_2, f = \lambda T.T[w := x]\} \uplus \\ &\quad \{(\{I_1 \mapsto 1, I_2 \mapsto 1\}, r_n, id, \\ &\quad \{(s(\tau) \cup U'_1) \mathbb{Q}_y^x \mapsto 1, s(\rho) \mathbb{Q}_w^y \mapsto 1, (s(\sigma) \cup U'_2) \mathbb{Q}_z^w \mapsto 1\}) \mid I_2 \in dom(S), \\ &\quad (I_1 = U_1 \mathbb{Q}_y^x, I_2 = U_2 \mathbb{Q}_z^x) \vee (I_2 = U_1 \mathbb{Q}_y^x, I_1 = U_2 \mathbb{Q}_z^x), \\ &\quad U_1 = \{\otimes_n^{\perp}(\rho).\tau\} \cup U'_1, U_2 = \{\otimes_n.\sigma\} \cup U'_2, w \notin E\} \end{aligned}$$

In the case of Brane Calculus, the unary reactions are only those arising from pinocytosis, while binary reactions arise from exocytosis and phagocytosis. In both cases, the multiplicity of each reactant is 1, so the multiplicity of  $I_1$  is not relevant. Exocytosis merges two compartments; this is reflected by the fact that the “rearranging” function  $f$  substitutes every occurrence of the name  $w$  in  $T$  with  $x$ . On the other hand, pinocytosis and phagocytosis create new compartments; to represent the new structure, we choose a fresh name  $w$  representing the new intermediate nesting level, and reconnect the various compartments accordingly. Therefore, for any reaction  $(S_1, r, f, S_2) \in reactions_E(I_1 \mapsto i, S)$ ,  $fn(S_2) \setminus E$  is either  $\emptyset$  (in the case of exocytosis) or the singleton  $\{w\}$ .

## 5.2 Adequacy results

Before proving the correctness of our implementation, we have to define how to translate a species set back to a system of the brane calculus.

Let  $S$  be a non empty species set. A *root name* of  $S$ , denoted by  $root(S)$ , is a name  $x$  such that  $S(C \mathbb{Q}_y^x) > 0$  for some  $C, y$ , and for all  $z, C' : C' \mathbb{Q}_z^x \notin dom(S)$ . The next result states that  $root(\_)$  is well defined on the species sets we encounter during a simulation.

**Lemma 5.2.** *For all  $P \in \mathbb{B}_{sys}$ :*

1. *if  $P \neq \diamond$ :  $root((|P|)_x) = x$ .*
2. *if  $(|P|)_x \xrightarrow{a, O} E \vdash (t, S, R)$  and  $S \neq \emptyset$ , then  $root(S) = x$ .*

*Proof.* (1.) is trivial by definition. (2.) It is enough to check that the reaction rules introduced by  $reactions(\_)$  do not change the name of the root, nor introduce new ones.  $\square$

We can now define a function  $\llbracket - \rrbracket$  which maps complexes to membranes, and a function  $\llbracket - \rrbracket_x$  mapping species sets to systems; the latter is parametric in the name  $x$  of the root of the system.

$$\llbracket A_1, \dots, A_n \rrbracket \triangleq A_1 | \dots | A_n \quad \llbracket S \rrbracket_x \triangleq \prod_{C \mathcal{C}_y^x \mathcal{D} \in \text{dom}(S)} S(C \mathcal{C}_y^x \mathcal{D}) \cdot (\llbracket C \rrbracket \mathcal{C} \llbracket S \rrbracket_y \mathcal{D})$$

where the notation  $n \cdot P$  is a shorthand for  $P \circ \dots \circ P$ ,  $n$  times.

**Lemma 5.3.** *For all  $P \in \mathbb{B}_{\text{sys}}$ :*

1.  $\llbracket \langle P \rangle_x \rrbracket_x \equiv P$ .
2. *if  $\langle P \rangle_x \xrightarrow{a, O} E \vdash (t, S, R)$  then  $\llbracket S \rrbracket_x$  is well defined.*

*Proof.* (1.) is easy. (2.) It is enough to check that the reaction rules introduced by  $\text{reactions}(-)$  do not introduce loops (i.e., the order among names is well founded).  $\square$

We can now state and prove the main results of this section.

**Proposition 5.4 (Soundness).** *For all  $P \in \mathbb{B}_{\text{sys}}$ , if  $\langle P \rangle_x \xrightarrow{a, O} E \vdash (t, S, R)$  then there exists  $\mu$  such that  $P \rightarrow \mu$  and  $\mu_{id}(\llbracket S \rrbracket_x) = a$ .*

*Proof.* The proof is by cases on which reaction rule  $O$  is selected by the function  $\text{next}$ . By additivity of measures, we can restrict ourselves to when the whole process  $P$  is the redex of the reduction. Let us see here the case when  $P$  is the redex of a (red-pino) (another is in Appendix B).

Let  $P = \otimes_n(\rho) \cdot \sigma | \sigma_0 \mathcal{C} P' \mathcal{D}$  and let us assume that  $\sigma_0$  does not exhibit a  $\otimes_n$  action. Then, the translation of  $P$  is  $\langle P \rangle_x = E \vdash \text{species}_{\emptyset, x}(P) \oplus (0, \emptyset, \emptyset)$ , where

$$\begin{aligned} \text{species}_{\emptyset, x}(P) &= \{s(\otimes_n(\rho) \cdot \sigma | \sigma_0) \mathcal{C}_y^x \mathcal{D} \mapsto 1\} \cup \text{species}_{\{y\}, x}(P') \\ &= \{(s(\otimes_n(\rho) \cdot \sigma) \cup s(\sigma_0)) \mathcal{C}_y^x \mathcal{D} \mapsto 1\} \cup \text{species}_{\{y\}, x}(P') \\ &= \{(\{\otimes_n(\rho) \cdot \sigma\} \cup s(\sigma_0)) \mathcal{C}_y^x \mathcal{D} \mapsto 1\} \cup \text{species}_{\{y\}, x}(P') \end{aligned}$$

Let  $I_1 = (\{\otimes_n(\rho) \cdot \sigma\} \cup s(\sigma_0)) \mathcal{C}_y^x \mathcal{D}$  and  $J_{P'} = \text{species}_{\{y\}, x}(P')$ ; then  $\langle P \rangle_x = I_1 \mapsto 1 \oplus J_{P'} \oplus (0, \emptyset, \emptyset) = J_{P'} \oplus (0, S', R')$  where

$$\begin{aligned} S' &= \{I_1 \mapsto 1\} \\ R' &= \text{init}(L, (0, S', \emptyset)) = \{O_L \mapsto (t_1, a_1)\} \\ O_L &= (\{I_1 \mapsto 1\}, r_n, id, \{(s(\sigma) \cup s(\sigma_0)) \mathcal{C}_y^x \mathcal{D} \mapsto 1, s(\rho) \mathcal{C}_w^y \mathcal{D} \mapsto 1\}) \\ L &= \text{reactions}(I_1 \mapsto 1, \emptyset) = \text{unary}(I_1) = \{(\{I_1 \mapsto 1\}, r_n, id, \{(s(\sigma) \cup s(\sigma_0)) \mathcal{C}_y^x \mathcal{D} \mapsto 1, s(\rho) \mathcal{C}_w^y \mathcal{D} \mapsto 1\})\} \end{aligned}$$

Now, the reaction  $O$  in  $\langle P \rangle_x \xrightarrow{a, O} T$  is  $O_L$  (otherwise it would involve  $P'$ , not the pino of the whole  $P$ ). This means that  $\langle P \rangle_x \xrightarrow{a, O} T$  is derived by means of an application of the (Reaction rule) as follows, where  $S_1 = \{I_1 \mapsto 1\}$  and  $S_2 = \{(s(\sigma) \cup s(\sigma_0)) \mathcal{C}_y^x \mathcal{D} \mapsto 1, s(\rho) \mathcal{C}_w^y \mathcal{D} \mapsto 1\}$ :

$$\frac{((S_1, r_n, id, S_2), a_1, t_1) = \text{next}(0, S', R') \quad (E' \vdash (0, S'', R'')) = \text{cow}(E \vdash (0, S', R'), S_1)}{E \vdash \{J_{P'}\} \oplus (0, S', R') \xrightarrow{a_1, (S_1, r_n, id, S_2)} \text{norm}(E \cup \text{fn}(S_2) \vdash (\{J_{P'}\} \oplus S_2 \oplus ((t_1, S'', R'') \ominus S_1)))}$$

where  $S'' = S'$  and  $R'' = R'$ .

$$\begin{aligned}
\{J_{P'}\} \oplus S_2 \oplus ((t_1, S'', R'') \oplus S_1) &= \\
&= J_{P'} \oplus \{(s(\sigma) \cup s(\sigma_0)) \mathcal{C}_y^x \mathcal{D} \mapsto 1, s(\rho) \mathcal{C}_w^y \mathcal{D} \mapsto 1\} \oplus ((t_1, S', R') \oplus \{I_1 \mapsto 1\}) \\
&= J_{P'} \oplus \{(s(\sigma) \cup s(\sigma_0)) \mathcal{C}_y^x \mathcal{D} \mapsto 1, s(\rho) \mathcal{C}_w^y \mathcal{D} \mapsto 1\} \oplus (t_1, \{I_1 \mapsto 0\}, \{O_L \mapsto (t_2, a_2)\}) \\
&= J_{P'} \oplus J' \oplus (t_1, \{I_1 \mapsto 0\}, \{O_L \mapsto (t_2, a_2)\})
\end{aligned}$$

Now let us define  $Q$  as  $Q = \sigma | \sigma_0 \mathcal{C} \rho \mathcal{C} \diamond \mathcal{D} \circ P' \mathcal{D}$ , then  $([Q])_x = \text{species}_{\emptyset, x}(Q) \oplus (0, \emptyset, \emptyset)$  where

$$\begin{aligned}
\text{species}_{\emptyset, x}(Q) &= \text{species}_{\emptyset, x}(\sigma | \sigma_0 \mathcal{C} \rho \mathcal{C} \diamond \mathcal{D} \circ P' \mathcal{D}) \\
&= \{s(\sigma | \sigma_0) \mathcal{C}_y^x \mathcal{D} \mapsto 1\} \cup \text{species}_{\{x\}, y}(\rho \mathcal{C} \diamond \mathcal{D} \circ P') \\
&= \{(s(\sigma) \cup s(\sigma_0)) \mathcal{C}_y^x \mathcal{D} \mapsto 1\} \cup \text{species}_{\{x\}, y}(\rho \mathcal{C} \diamond \mathcal{D}) \cup \text{species}_{\{x\}, y}(P') \\
&= \{(s(\sigma) \cup s(\sigma_0)) \mathcal{C}_y^x \mathcal{D} \mapsto 1\} \cup \{s(\rho) \mathcal{C}_w^y \mathcal{D} \mapsto 1\} \cup \emptyset \cup J_{P'} = J' \cup J_{P'}
\end{aligned}$$

and hence  $Q \equiv [[J_{P'} \oplus J' \oplus (t_1, \{I_1 \mapsto 0\}, \{O_L \mapsto (t_2, a_2)\})]]_x$ . It remains to prove that  $r_n = \mu_{id}([Q])$ . Let us notice that the derivation of  $P \rightarrow \mu$  is actually as follows:

$$\begin{array}{c}
\text{(par)} \frac{\textcircled{n}(\rho) \cdot \sigma \rightarrow [\textcircled{n}]_{\sigma}^{\rho} \quad \sigma_0 \rightarrow \mu''}{\textcircled{n}(\rho) \cdot \sigma | \sigma_0 \rightarrow [\textcircled{n}]_{\sigma \textcircled{n}(\rho) \cdot \sigma \oplus \sigma_0}^{\rho} \mu''} \quad P' \rightarrow \mu' \\
\text{(loc)} \frac{}{\textcircled{n}(\rho) \cdot \sigma | \sigma_0 \mathcal{C} P' \mathcal{D} \rightarrow \mu}
\end{array}$$

where  $\mu = \mu' @_{P'}^{\textcircled{n}(\rho) \cdot \sigma | \sigma_0}([\textcircled{n}]_{\sigma \textcircled{n}(\rho) \cdot \sigma \oplus \sigma_0}^{\rho} \mu'')$ . Then:

$$\begin{aligned}
\mu_{id}([\sigma | \sigma_0 \mathcal{C} \rho \mathcal{C} \diamond \mathcal{D} \circ P' \mathcal{D}]) &= (\mu' @_{P'}^{\textcircled{n}(\rho) \cdot \sigma | \sigma_0}([\textcircled{n}]_{\sigma \textcircled{n}(\rho) \cdot \sigma \oplus \sigma_0}^{\rho} \mu''))_{id}([\sigma | \sigma_0 \mathcal{C} \rho \mathcal{C} \diamond \mathcal{D} \circ P' \mathcal{D}]) \\
&= \mu'_{id}([\sigma | \sigma_0 \mathcal{C} \rho \mathcal{C} \diamond \mathcal{D} \circ P' \mathcal{D}]) + ([\textcircled{n}]_{\sigma \textcircled{n}(\rho) \cdot \sigma \oplus \sigma_0}^{\rho} \mu'')_{\textcircled{n}}([\sigma | \sigma_0] \times [\rho]) \\
&= ([\textcircled{n}]_{\sigma \textcircled{n}(\rho) \cdot \sigma \oplus \sigma_0}^{\rho} \mu'')_{\textcircled{n}}([\sigma | \sigma_0] \times [\rho]) \\
&= ([\textcircled{n}]_{\sigma}^{\rho})_{\textcircled{n}}([\sigma | \sigma_0] \times [\rho]) + \mu''_{\textcircled{n}}([\sigma | \sigma_0] \times [\rho]) \\
&= r_n + \mu''_{\textcircled{n}}([\sigma | \sigma_0] \times [\rho]) = r_n
\end{aligned}$$

where the last equivalence holds because  $\mu''_{\textcircled{n}}([\sigma | \sigma_0] \times [\rho]) = 0$  because we assumed that the reaction does not involve  $\sigma_0$ .  $\square$

**Proposition 5.5 (Progress).** *For all processes  $P, Q$ , if  $P \twoheadrightarrow Q$  then there exists a reaction  $O$  and a term  $T$  such that  $([P])_x \xrightarrow{a, O} T$  and  $Q \equiv [[T]]_x$ .*

*Proof.* By induction on the derivation of  $P \twoheadrightarrow Q$ . Let us see the case of (red-pin), the others being similar. Let  $P = \textcircled{n}(\rho) \cdot \sigma | \sigma_0 \mathcal{C} P' \mathcal{D}$  and  $Q = \sigma | \sigma_0 \mathcal{C} \rho \mathcal{C} \diamond \mathcal{D} \circ P' \mathcal{D}$ . Then,

$$([P])_x = \text{species}_{\{y\}, x}(P') \oplus (0, \{(\textcircled{n}(\rho) \cdot \sigma) \cup s(\sigma_0)\} \mathcal{C}_y^x \mathcal{D} \mapsto 1, \{O_L \mapsto (t_1, a_1)\})$$

where  $O_L = (\{(\textcircled{n}(\rho) \cdot \sigma) \cup s(\sigma_0)\} \mathcal{C}_y^x \mathcal{D} \mapsto 1, r_n, id, \{(s(\sigma) \cup s(\sigma_0)) \mathcal{C}_y^x \mathcal{D} \mapsto 1, s(\rho) \mathcal{C}_w^y \mathcal{D} \mapsto 1\})$ . Then, by the (Reaction rule) we can take  $T = \text{species}_{\{x\}, y}(P') \oplus (t_1, \{(s(\sigma) \cup s(\sigma_0)) \mathcal{C}_y^x \mathcal{D} \mapsto 1, s(\rho) \mathcal{C}_w^y \mathcal{D} \mapsto 1\}, \emptyset)$ . It is easy to check that  $Q \equiv [[T]]_x$ .  $\square$

**Proposition 5.6 (Completeness).** *For all processes  $P, Q$ , if  $P \rightarrow \mu$  and  $\mu_{id}([Q]) > 0$  then for all node name  $x$ , there exists a reaction  $O$  and a term  $T$  such that  $([P])_x \xrightarrow{a, O} T$ ,  $Q \equiv [[T]]_x$  and  $a = \mu_{id}([Q])$ .*

*Proof.* If  $P \rightarrow \mu$  and  $\mu_{id}([Q]) > 0$  then  $P \twoheadrightarrow Q$  by Prop. 3.4. By Prop. 5.5, we have that for some  $a, O, T$ ,  $([P])_x \xrightarrow{a, O} T$  and  $Q \equiv [[T]]_x$ . But then  $a = \mu_{id}([Q])$  by soundness (Prop. 5.4).  $\square$

### 5.3 Example

We conclude this section with an example. Let  $P = 10000 \cdot \mathfrak{v}_n \cdot \mathfrak{v}_m \langle \mathfrak{v}_k \rangle \circ 100 \cdot ((\mathfrak{v}_n^\perp(\mathfrak{v}_m^\perp) \mid \mathfrak{v}^\perp) \langle \mathfrak{v}_k \rangle)$ . Then, its reductions in the Brane Calculus are as follows:

$$\begin{aligned} & 10000 \cdot \mathfrak{v}_n \cdot \mathfrak{v}_m \langle \mathfrak{v}_k \rangle \circ 100 \cdot ((\mathfrak{v}_n^\perp(\mathfrak{v}_m^\perp) \mid \mathfrak{v}^\perp) \langle \mathfrak{v}_k \rangle) \\ \longrightarrow & 9999 \cdot \mathfrak{v}_n \cdot \mathfrak{v}_m \langle \mathfrak{v}_k \rangle \circ 99 \cdot ((\mathfrak{v}_n^\perp(\mathfrak{v}_m^\perp) \mid \mathfrak{v}^\perp) \langle \mathfrak{v}_k \rangle) \circ \mathfrak{v}^\perp \langle \mathfrak{v}_m^\perp \rangle \langle \mathfrak{v}_m \rangle \langle \mathfrak{v}_k \rangle \circ \mathfrak{v}_k \langle \mathfrak{v} \rangle \\ \longrightarrow & 9999 \cdot \mathfrak{v}_n \cdot \mathfrak{v}_m \langle \mathfrak{v}_k \rangle \circ 99 \cdot ((\mathfrak{v}_n^\perp(\mathfrak{v}_m^\perp) \mid \mathfrak{v}^\perp) \langle \mathfrak{v}_k^\perp \rangle) \circ \mathfrak{v}^\perp \langle \mathfrak{v}_k \rangle \circ \diamond \langle \mathfrak{v} \rangle \end{aligned}$$

The translation of  $P$  is  $\langle P \rangle_x = E \vdash \text{species}_{\emptyset, x}(P) \oplus (0, \emptyset, \emptyset)$ , where

$$\begin{aligned} \text{species}_{\emptyset, x}(P) &= \text{species}_{\{x\}}(10000 \cdot \mathfrak{v}_n \cdot \mathfrak{v}_m \langle \mathfrak{v}_k \rangle \circ 100 \cdot (\mathfrak{v}_n^\perp(\mathfrak{v}_m^\perp) \mid \mathfrak{v}^\perp) \langle \mathfrak{v}_k \rangle) \\ &= \{s(\mathfrak{v}_n \cdot \mathfrak{v}_m) \langle \mathfrak{v}_y \rangle \mapsto 10000\} \cup \text{species}_{\{x, y\}, y}(\mathfrak{v}_k \langle \mathfrak{v} \rangle) \cup \{s(\mathfrak{v}_n^\perp(\mathfrak{v}_m^\perp) \mid \mathfrak{v}^\perp) \langle \mathfrak{v}_z \rangle \mapsto 100\} \\ &\quad \cup \text{species}_{\{x, z\}, z}(\mathfrak{v}_k \langle \mathfrak{v} \rangle) \\ &= \{\{\mathfrak{v}_n \cdot \mathfrak{v}_m\} \langle \mathfrak{v}_y \rangle \mapsto 10000, \{\mathfrak{v}_n^\perp(\mathfrak{v}_m^\perp), \mathfrak{v}^\perp\} \langle \mathfrak{v}_z \rangle \mapsto 100, \{\mathfrak{v}_k\} \langle \mathfrak{v}_w \rangle \mapsto 1, \\ &\quad \{\mathfrak{v}_k\} \langle \mathfrak{v}_v \rangle \mapsto 1\} \end{aligned}$$

Let  $I_1 = \{\mathfrak{v}_n \cdot \mathfrak{v}_m\} \langle \mathfrak{v}_y \rangle$ ,  $I_2 = \{\mathfrak{v}_n^\perp(\mathfrak{v}_m^\perp), \mathfrak{v}^\perp\} \langle \mathfrak{v}_z \rangle$ ,  $I_3 = \{\mathfrak{v}_k\} \langle \mathfrak{v}_w \rangle$ ,  $I_4 = \{\mathfrak{v}_k\} \langle \mathfrak{v}_v \rangle$ ,  $r_n = 10s^{-1}$ ,  $r_k = 5s^{-1}$  and  $r_m = 5s^{-1}$ ; then

$$\begin{aligned} \langle P \rangle_x &= I_1 \mapsto 10000 \oplus I_2 \mapsto 100 \oplus I_3 \mapsto 1 \oplus I_4 \mapsto 1 \oplus (0, \emptyset, \emptyset) \\ &= I_2 \mapsto 100 \oplus I_3 \mapsto 1 \oplus I_4 \mapsto 1 \oplus (0, S_1, R_1) \\ &= I_3 \mapsto 1 \oplus I_4 \mapsto 1 \oplus (0, S_2, R_2) \\ &= I_4 \mapsto 1 \oplus (0, S_3, R_3) = (0, S_4, R_4) \end{aligned}$$

where

$$\begin{aligned} L_1 &= \text{reactions}(I_1 \mapsto 10000, \emptyset) = \emptyset \\ S_1 &= S\{I_1 \mapsto 10000\} \\ R_1 &= R \cup \text{init}(L_1, (0, S_1, \emptyset)) = \emptyset \\ L_2 &= \text{reactions}(I_2 \mapsto 100, S_1) \\ &= (\{I_2 \mapsto 1, I_1 \mapsto 1\}, 10, \text{id}, \{\{\mathfrak{v}^\perp\} \langle \mathfrak{v}_z \rangle \mapsto 1, \{\mathfrak{v}_m^\perp\} \langle \mathfrak{v}_w \rangle \mapsto 1, \{\mathfrak{v}_m\} \langle \mathfrak{v}_y \rangle \mapsto 1\}) \\ S_2 &= S_1 \{I_2 \mapsto 100\} = \{I_1 \mapsto 10000, I_2 \mapsto 100\} \\ R_2 &= R_1 \cup \text{init}(L_2, (0, S_2, R_1)) = \{O_{L_2} \mapsto (t_1, a_1)\} \\ L_3 &= \text{reactions}(I_3 \mapsto 1, S_2) = \emptyset \\ S_3 &= S_2 \{I_3 \mapsto 1\} = \{I_1 \mapsto 10000, I_2 \mapsto 100, I_3 \mapsto 1\} \\ R_3 &= R_2 \cup \text{init}(L_3, (0, S_3, R_2)) = R_2 \\ L_4 &= \text{reactions}(I_4 \mapsto 1, S_3) = \emptyset \\ S_4 &= S_3 \{I_4 \mapsto 1\} = \{I_1 \mapsto 10000, I_2 \mapsto 100, I_3 \mapsto 1, I_4 \mapsto 1\} \\ R_4 &= R_3 \cup \text{init}(L_4, (0, S_4, R_3)) = R_3 \end{aligned}$$

with  $O_{L_2} = (\{I_2 \mapsto 1, I_1 \mapsto 1\}, 10, id, \{\{\vartheta^\perp\} \mathcal{C}_z^x \mapsto 1, \{\vartheta_m^\perp\} \mathcal{C}_w^z \mapsto 1, \{\vartheta_m\} \mathcal{C}_y^w \mapsto 1\})$   
 $a_1 = propensity(O_{L_2}, S_2) = 10000000$   
 $t_1 = 0 + delay(10, 10000000).$

Now, the reaction  $O$  in  $(\mathcal{P})_x \xrightarrow{a, O} T$  is  $O_L$ . This means that  $(\mathcal{P})_x \xrightarrow{a, O} T$  is derived by means of an application of the (Reaction rule) as follows, where  $E = x, y, z, w$ ,  $S_1 = \{I_2 \mapsto 1, I_1 \mapsto 1\}$ , and  $S_2 = \{\{\vartheta^\perp\} \mathcal{C}_z^x \mapsto 1, \{\vartheta_m^\perp\} \mathcal{C}_w^z \mapsto 1, \{\vartheta_m\} \mathcal{C}_y^w \mapsto 1\}$ :

$$\frac{((S_1, 10, id, S_2), 10, t_1) = next(0, S_4, R_4) \quad (E' \vdash (0, S_5, R_5)) = cow(E \vdash (0, S_4, R_4), S_1)}{E \vdash (0, S_4, R_4) \xrightarrow{10, (S_1, 10, id, S_2)} norm(E' \cup fn(S_2) \vdash (S_2 \oplus ((t_1, S_5, R_5) \ominus S_1)))}$$

where

$$S_5 = \{I_1 \mapsto 1, \{\vartheta_n, \vartheta_m\} \mathcal{C}_y^x \mapsto 9999, I_2 \mapsto 1, \{\vartheta_n^\perp(\vartheta_m^\perp), \vartheta^\perp\} \mathcal{C}_z^x \mapsto 99, I_3 \mapsto 1, \\ I_4 \mapsto 1, \{\vartheta_k\} \mathcal{C}_w^y \mapsto 1, \{\vartheta_k\} \mathcal{C}_v^z \mapsto 1\}$$

$$R_5 = \{O_{L_2} \mapsto (t_1, a_1), O_1 \mapsto (t_2, a_2), O_2 \mapsto (t_3, a_3), O_3 \mapsto (t_4, a_4)\}$$

with

$$O_1 = (\{\{\vartheta_n, \vartheta_m\} \mathcal{C}_y^x \mapsto 1, I_2 \mapsto 1\}, 10, id, \{\{\vartheta^\perp\} \mathcal{C}_z^x \mapsto 1, \{\vartheta_m^\perp\} \mathcal{C}_w^z \mapsto 1, \{\vartheta_m\} \mathcal{C}_y^w \mapsto 1\})$$

$$O_2 = (\{\vartheta_n^\perp(\vartheta_m^\perp), \vartheta^\perp\} \mathcal{C}_z^x \mapsto 1, I_1 \mapsto 1\}, 10, id, \{\{\vartheta^\perp\} \mathcal{C}_z^x \mapsto 1, \{\vartheta_m^\perp\} \mathcal{C}_w^z \mapsto 1, \{\vartheta_m\} \mathcal{C}_y^w \mapsto 1\})$$

$$O_3 = (\{\{\vartheta_n, \vartheta_m\} \mathcal{C}_y^x \mapsto 1, \{\vartheta_n^\perp(\vartheta_m^\perp), \vartheta^\perp\} \mathcal{C}_z^x \mapsto 1\}, 10, id, \{\{\vartheta^\perp\} \mathcal{C}_z^x \mapsto 1, \{\vartheta_m^\perp\} \mathcal{C}_w^z \mapsto 1, \{\vartheta_m\} \mathcal{C}_y^w \mapsto 1\})$$

$$a_2 = propensity(O_1, S_5) = 99990$$

$$a_3 = propensity(O_2, S_5) = 990$$

$$a_4 = propensity(O_3, S_5) = 989901$$

$$t_2 = 0 + delay(10, 99990)$$

$$t_3 = 0 + delay(10, 990)$$

$$t_4 = 0 + delay(10, 989901)$$

We can now compute the multiset of the new machine state:

$$\begin{aligned} S_2 \oplus ((t_1, S_5, R_5) \ominus S_1) &= S_2 \oplus ((t_1, S_5, R_5) \ominus \{I_2 \mapsto 1, I_1 \mapsto 1\}) \\ &= S_2 \oplus ((t_1, S_6, R_6) \ominus \{I_1 \mapsto 1\}) \\ &= \{\{\vartheta^\perp\} \mathcal{C}_z^x \mapsto 1, \{\vartheta_m^\perp\} \mathcal{C}_w^z \mapsto 1, \{\vartheta_m\} \mathcal{C}_y^w \mapsto 1\} \oplus (t_1, S_7, R_7) \\ &= (t_1, S_8, R_8) \end{aligned}$$

with

$$\begin{aligned} a'_2 &= propensity(O_1, S_6) = 0 & S_8 &= S_6 \cup \{\{\vartheta^\perp\} \mathcal{C}_z^x \mapsto 1, \{\vartheta_m^\perp\} \mathcal{C}_w^z \mapsto 1, \{\vartheta_m\} \mathcal{C}_y^w \mapsto 1\} \\ t'_2 &= t_1 + (a_2/a'_2)(t_2 - t_1) & R_6 &= \{O_1 \mapsto (t'_2, a'_2)\} \\ a'_3 &= propensity(O_2, S_7) = 0 & R_7 &= R_6 \cup \{O_2 \mapsto (t'_3, a'_3)\} \\ t'_3 &= t_1 + (a_3/a_5)(t_3 - t_1) & R_8 &= R_7 \cup \{\{\vartheta_m^\perp\} \mathcal{C}_w^z \mapsto 1, \{\vartheta_m\} \mathcal{C}_y^w \mapsto 1\}, 5, f, \{\emptyset \mathcal{C}_w^z \mapsto 1\} \\ S_6 &= S_5 \setminus \{I_2 \mapsto 1\} \\ S_7 &= S_6 \setminus \{I_1 \mapsto 1\} & f &= \lambda T. T[y := z] \end{aligned}$$

## 6 Conclusions

In this paper, we have presented an abstract machine for the Stochastic Brane Calculus. Instead of defining an *ad hoc* machine, we have adopted the generic abstract machine for stochastic calculi (GSAM) recently introduced by Lakin, Paulevé and Phillips. According to the encoding technique we have adopted, membranes are flattened into a set of species, where the hierarchical structure is represented by means of *names*. In order to keep track of these names, and for dealing efficiently with multiple copies of the same species, we have introduced a new generic abstract machine, called COWGSAM, which extends the GSAM with a *name environment* and a *copy-on-write* optimization strategy. We have proved that the implementation of the Stochastic Brane Calculus in COWGSAM is adequate with respect to the stochastic structural operational semantics of the calculus given in [2].

We think that COWGSAM can be used for implementing other stochastic calculi dealing with nested structures, also beyond the models for systems biology. In particular, it is interesting to apply this approach to Stochastic Bigraphs [9], a general meta-model well-suited for representing a range of stochastic systems with compartments; in this way we would obtain a *General Stochastic Bigraphical Machine*, which could be instantiated to any given stochastic bigraphic reactive system. However, such a machine would not scale well, as in general the COW strategy may be not very useful; thus, we can restrict our attentions to smaller subsets of BRSs, specifically designed to some application domain. For biological applications, the bigraphic reactive systems considered in [1, 7] might be a more reasonable target.

Another interesting question is about the expressive power of GSAM and COWGSAM. We think that GSAM correspond to stochastic (multiset) Petri nets, but COWGSAM could go further thanks to the possibility of creating unlimited new names during execution. Further work include comparison with other stochastic simulation tools dealing with compartments, like BioPEPA [6].

**Acknowledgment** Work funded by MIUR PRIN project “SisteR”, prot. 20088HXMYN.

## References

- [1] Giorgio Bacci, Davide Grohmann & Marino Miculan (2009): *Bigraphical models for protein and membrane interactions*. In: Gabriel Ciobanu, editor: *Proc. MeCBIC’09, EPTCS 11*, pp. 3–18. doi:10.4204/EPTCS.11.1.
- [2] Giorgio Bacci & Marino Miculan (2012): *Measurable Stochastics for Brane Calculus*. *Theoretical Computer Science* 431, pp. 117–136. doi:10.1016/j.tcs.2011.12.055.
- [3] Roberto Barbuti, Andrea Maggiolo-Schettini, Paolo Milazzo & Angelo Troina (2007): *The Calculus of Looping Sequences for Modeling Biological Membranes*. In: George Eleftherakis, Petros Kefalas, Gheorghe Paun, Grzegorz Rozenberg & Arto Salomaa, editors: *Workshop on Membrane Computing, Lecture Notes in Computer Science* 4860, Springer, pp. 54–76. doi:10.1007/978-3-540-77312-2\_4.
- [4] Luca Cardelli (2004): *Brane Calculi*. In: Vincent Danos & Vincent Schächter, editors: *Proc. CMSB, Lecture Notes in Computer Science* 3082, Springer, pp. 257–278. doi:10.1007/978-3-540-25974-9\_24.
- [5] Luca Cardelli & Radu Mardare (2010): *The Measurable Space of Stochastic Processes*. In: *Proc. QEST*, IEEE Computer Society, pp. 171–180. doi:10.1109/QEST.2010.30.
- [6] Federica Ciocchetta & Maria Luisa Guerriero (2009): *Modelling Biological Compartments in Bio-PEPA*. *Electronic Notes in Theoretical Computer Science* 227, pp. 77–95. doi:10.1016/j.entcs.2008.12.105.
- [7] Troels Christoffer Damgaard, Espen Højsgaard & Jean Krivine (2012): *Formal Cellular Machinery*. *Electronic Notes in Theoretical Computer Science* 284, pp. 55–74. doi:10.1016/j.entcs.2012.05.015.
- [8] Holger Hermanns (2002): *Interactive Markov Chains: The Quest for Quantified Quality*, *Lecture Notes in Computer Science* 2428. Springer. doi:10.1007/3-540-45804-2.

- [9] Jean Krivine, Robin Milner & Angelo Troina (2008): *Stochastic Bigraphs*. In: *Proc. 24th MFPS, ENTCS* 218, pp. 73–96. doi:10.1016/j.entcs.2008.10.006.
- [10] Matthew R. Lakin, Loïc Paulevé & Andrew Phillips (2012): *Stochastic Simulation of Multiple Process Calculi for Biology*. *Theoretical Computer Science* 431, pp. 181–206. doi:10.1016/j.tcs.2011.12.057.
- [11] Cosimo Laneve & Fabien Tarissan (2008): *A simple calculus for proteins and cells*. *Theoretical Computer Science* 404(1-2), pp. 127–141. doi:10.1016/j.tcs.2008.04.011.
- [12] Prakash Panangaden (2009): *Labelled Markov Processes*. Imperial College Press, London, U.K.
- [13] Loïc Paulevé, Simon Youssef, Matthew R. Lakin & Andrew Phillips (2010): *A generic abstract machine for stochastic process calculi*. In: Paola Quaglia, editor: *Proc. CMSB*, ACM, pp. 43–54. doi:10.1145/1839764.1839771.
- [14] Aviv Regev, Ekaterina M. Panina, William Silverman, Luca Cardelli & Ehud Y. Shapiro (2004): *BioAmbients: an abstraction for biological compartments*. *Theoretical Computer Science* 325(1), pp. 141–167. doi:10.1016/j.tcs.2004.03.061.

## A Some measure theory

Given a set  $M$ , a family  $\Sigma$  of subsets of  $M$  is called a  $\sigma$ -algebra if it contains  $M$  and is closed under complements and (infinite) countable unions:

1.  $M \in \Sigma$ ;
2.  $A \in \Sigma$  implies  $A^c \in \Sigma$ , where  $A^c = M \setminus A$ ;
3.  $\{A_i\}_{i \in \mathbb{N}} \subset \Sigma$  implies  $\bigcup_{i \in \mathbb{N}} A_i \in \Sigma$ .

Since  $M \in \Sigma$  and  $M^c = \emptyset$ ,  $\emptyset \in \Sigma$ , hence  $\Sigma$  is nonempty by definition. A  $\sigma$ -algebra is closed under countable set-theoretic operations: is closed under finite unions ( $A, B \in \Sigma$  implies  $A \cup B = A \cup B \cup \emptyset \cup \emptyset \cup \dots \in \Sigma$ ), countable intersections (by DeMorgan's law  $A \cap B = (A^c \cup B^c)^c$  in its finite and infinite version), and countable subtractions ( $A, B \in \Sigma$  implies  $A \setminus B = A \cap B^c \in \Sigma$ ).

**Definition A.1 (Measurable Space).** *Given a set  $M$  and a  $\sigma$ -algebra on  $M$ , the tuple  $(M, \Sigma)$  is called a measurable space, the elements of  $\Sigma$  measurable sets, and  $M$  the support-set.*

A set  $\Omega \subseteq 2^M$  is a *generator for the  $\sigma$ -algebra  $\Sigma$  on  $M$*  if  $\Sigma$  is the closure of  $\Omega$  under complement and countable union; we write  $\sigma(\Omega) = \Sigma$  and say that  $\Sigma$  is generated by  $\Omega$ . Note that the  $\sigma$ -algebra generated by a  $\Omega$  is also the smallest  $\sigma$ -algebra containing  $\Omega$ , that is, the intersection of all  $\sigma$ -algebras that contain  $\Omega$ . In particular it holds that a completely arbitrary intersection of  $\sigma$ -algebras is a  $\sigma$ -algebra. A  $\sigma$ -algebra generated by  $\Omega$ , denoted by  $\sigma(\Omega)$ , is minimal in the sense that if  $\Omega \subset \Sigma$  and  $\Sigma$  is a  $\sigma$ -algebra, then  $\sigma(\Omega) \subset \Sigma$ . If  $\Omega$  is a  $\sigma$ -algebra then obviously  $\sigma(\Omega) = \Omega$ ; if  $\Omega$  is empty or  $\Omega = \{\emptyset\}$ , or  $\Omega = \{M\}$ , then  $\sigma(\Omega) = \{\emptyset, M\}$ ; if  $\Omega \subset \Sigma$  and  $\Sigma$  is a  $\sigma$ -algebra, then  $\sigma(\Omega) \subset \Sigma$ . A generator  $\Omega$  for  $\Sigma$  is a *base for  $\Sigma$*  if it has disjoint elements. Note that if  $\Omega$  is a base for  $\Sigma$ , all measurable sets in  $\Sigma$  can be decomposed into countable unions of elements in  $\Omega$ .

A *measure* on a measurable space  $(M, \Sigma)$  is a function  $\mu: \Sigma \rightarrow \mathbb{R}_{\infty}^+$ , where  $\mathbb{R}_{\infty}^+$  denotes the extended positive real line, such that

1.  $\mu(\emptyset) = 0$ ;
2. for any disjoint sequence  $\{N_i\}_{i \in I} \subseteq \Sigma$  with  $I \subseteq \mathbb{N}$ , it holds

$$\mu\left(\bigcup_{i \in I} N_i\right) = \sum_{i \in I} \mu(N_i).$$

The triple  $(M, \Sigma, \mu)$  is called a *measure space*. A measure space  $(M, \Sigma, \mu)$  is called *finite* if  $\mu(M)$  is a finite real number; it is called  *$\sigma$ -finite* if  $M$  can be decomposed into a countable union of measurable sets of finite measure, that is,  $M = \bigcup_{i \in I} N_i$ , for some  $I \subseteq \mathbb{N}$  and  $\mu(N_i) \in \mathbb{R}^+$  for each  $i \in I$ . A set in a measure space has  *$\sigma$ -finite measure* if it is a countable union of sets with finite measure. Specifying a measure includes specifying its domain. If  $\mu$  is a measure on a measurable space  $(M, \Sigma)$  and  $\Sigma'$  is a  $\sigma$ -algebra contained in  $\Sigma$ , then the restriction  $\mu'$  of  $\mu$  to  $\Sigma'$  is also a measure, and in particular a measure on  $(M', \Sigma')$ , for some  $M' \subseteq M$  such that  $\Sigma'$  is a  $\sigma$ -algebra on  $M'$ .

Given two measurable spaces and measures on them, one can obtain the product measurable space and the product measure on that space. Let  $(M_1, \Sigma_1)$  and  $(M_2, \Sigma_2)$  be measurable spaces, and  $\mu_1$  and  $\mu_2$  be measures on these spaces. Denote by  $\Sigma_1 \otimes \Sigma_2$  the  $\sigma$ -algebra on the cartesian product  $M_1 \times M_2$  generated by subsets of the form  $B_1 \times B_2$ , said rectangles, where  $B_1 \in \Sigma_1$  and  $B_2 \in \Sigma_2$ . The *product measure*  $\mu_1 \otimes \mu_2$  is defined to be the unique measure on the measurable space  $(M_1 \times M_2, \Sigma_1 \otimes \Sigma_2)$  such that, for all  $B_1 \in \Sigma_1$  and  $B_2 \in \Sigma_2$

$$(\mu_1 \otimes \mu_2)(B_1 \times B_2) = \mu_1(B_1) \cdot \mu_2(B_2)$$

The existence of this measure is guaranteed by the Hahn-Kolmogorov theorem. The uniqueness of the product measure is guaranteed only in the case that both  $(M_1, \Sigma_1, \mu_1)$  and  $(M_2, \Sigma_2, \mu_2)$  are  $\sigma$ -finite.

Let  $\Delta(M, \Sigma)$  be the family of measures on  $(M, \Sigma)$ . It can be organized as a measurable space by considering the  $\sigma$ -algebra generated by the sets  $\{\mu \in \Delta(M, \Sigma) : \mu(S) \geq r\}$ , for arbitrary  $S \in \Sigma$  and  $r > 0$ .

Given two measurable spaces  $(M, \Sigma)$  and  $(N, \Theta)$  a mapping  $f: M \rightarrow N$  is *measurable* if for any  $T \in \Theta$ ,  $f^{-1}(T) \in \Sigma$ . Measurable functions are closed under composition: given  $f: M \rightarrow N$  and  $g: N \rightarrow O$  measurable functions then  $g \circ f: M \rightarrow O$  is also measurable.

## B Proof of Prop. 5.4

Let  $P = \mathfrak{v}_n^\perp \cdot \tau | \tau_0 \mathfrak{d} \mathfrak{v}_n \cdot \sigma | \sigma_0 \mathfrak{d} \mathfrak{P}' \mathfrak{d} \circ \mathfrak{P}'' \mathfrak{d}$ ; then,  $(\mathfrak{P})_x = E \vdash \text{species}_{\emptyset, x}(P) \oplus (0, \emptyset, \emptyset)$ , where

$$\begin{aligned} \text{species}_{\emptyset, x}(P) &= \{s(\mathfrak{v}_n^\perp \cdot \tau | \tau_0) \mathfrak{d}_y^x \mathfrak{d}\} \cup \text{species}_{\{x\}, y}(\mathfrak{v}_n \cdot \sigma | \sigma_0 \mathfrak{d} \mathfrak{P}' \mathfrak{d} \circ \mathfrak{P}'' \mathfrak{d}) \\ &= \{(\{\mathfrak{v}_n^\perp \cdot \tau\} \cup s(\tau_0)) \mathfrak{d}_y^x \mathfrak{d} \mapsto 1\} \cup \text{species}_{\{x\}, y}(\mathfrak{v}_n \cdot \sigma | \sigma_0 \mathfrak{d} \mathfrak{P}' \mathfrak{d}) \cup \text{species}_{\{x\}, y}(\mathfrak{P}'') \\ &= \{(\{\mathfrak{v}_n^\perp \cdot \tau\} \cup s(\tau_0)) \mathfrak{d}_y^x \mathfrak{d} \mapsto 1\} \cup \{s(\mathfrak{v}_n \cdot \sigma | \sigma_0) \mathfrak{d}_w^y \mathfrak{d} \mapsto 1\} \cup \text{species}_{\{x, y\}, w}(\mathfrak{P}') \cup \text{species}_{\{x\}, y}(\mathfrak{P}'') \\ &= \{(\{\mathfrak{v}_n^\perp \cdot \tau\} \cup s(\tau_0)) \mathfrak{d}_y^x \mathfrak{d} \mapsto 1\} \cup \{(s(\mathfrak{v}_n \cdot \sigma) \cup s(\sigma_0)) \mathfrak{d}_w^y \mathfrak{d} \mapsto 1\} \cup \text{species}_{\{x, y\}, w}(\mathfrak{P}') \cup \text{species}_{\{x\}, y}(\mathfrak{P}'') \\ &= \{(\{\mathfrak{v}_n^\perp \cdot \tau\} \cup s(\tau_0)) \mathfrak{d}_y^x \mathfrak{d} \mapsto 1\} \cup \{(\{\mathfrak{v}_n \cdot \sigma\} \cup s(\sigma_0)) \mathfrak{d}_w^y \mathfrak{d} \mapsto 1\} \cup \text{species}_{\{x, y\}, w}(\mathfrak{P}') \cup \text{species}_{\{x\}, y}(\mathfrak{P}'') \end{aligned}$$

Let  $I_1 = (\{\mathfrak{v}_n^\perp \cdot \tau\} \cup s(\tau_0)) \mathfrak{d}_y^x \mathfrak{d}$ ,  $I_2 = (\{\mathfrak{v}_n \cdot \sigma\} \cup s(\sigma_0)) \mathfrak{d}_w^y \mathfrak{d}$ ,  $J_{P'} = \text{species}_{\{x, y\}, w}(\mathfrak{P}')$ ,  $J_{P''} = \text{species}_{\{x\}, y}(\mathfrak{P}'')$ ; then  $(\mathfrak{P})_x = I_1 \mapsto 1 \oplus I_2 \mapsto 1 \oplus J_{P'} \oplus J_{P''} \oplus (0, \emptyset, \emptyset) = I_2 \mapsto 1 \oplus J_{P'} \oplus J_{P''} \oplus (0, S', R') = J_{P'} \oplus J_{P''} \oplus (0, S'', R'')$  where

$$\begin{aligned} L' &= \text{reactions}(I_1 \mapsto 1, \emptyset) = \emptyset \\ S' &= S\{I_1 \mapsto 1\} \\ R' &= \text{init}(L', (0, S', \emptyset)) = \emptyset \\ L'' &= \text{reactions}(I_2 \mapsto 1, S') \\ &= (\{I_2 \mapsto 1, I_1 \mapsto 1\}, r_n, f, \{(s(\tau) \cup s(\tau_0)) \cup s(\sigma) \cup s(\sigma_0)) \mathfrak{d}_y^x \mathfrak{d} \mapsto 1\}) \\ S'' &= S'\{I_2 \mapsto 1\} = \{I_1 \mapsto 1, I_2 \mapsto 1\} \\ R'' &= \text{init}(L'', (0, S'', R')) = \{O_L \mapsto (t_1, a_1)\} \end{aligned}$$

